

**Modal verbs *can, could, may, might, should* and *must* and semi-modals
ought to and *have to* in spoken Scottish English as compared to spoken
English English**

Suvi Myllyniemi
University of Tampere
School of Language, Translation and Literary Studies
English Language and Literature
MA Thesis
November 2015

Tampereen yliopisto
Kieli-, käännös- ja kirjallisuustieteiden yksikkö
Englannin kieli ja kirjallisuus

MYLLYNIEMI, SUVI: Modal verbs *can, could, may, might, should* and *must* and semi-modals *ought to* and *have to* in spoken Scottish English as compared to spoken English English

Pro gradu -tutkielma, 64 s.
Marraskuu 2015

Tämä pro gradu –tutkielma tarkastelee modaaliapuverbien *can, could, may, might, should* ja *must* sekä semimodaalien *ought to* ja *have to* käyttöä puhutussa skotti- ja englanninenglannissa vertaillen näitä keskenään siten että pääpaino on skottienglannissa. Tarkoituksena on selvittää, missä suhteessa kukin modaaliapuverbi tai semimodaali edustaa kutakin kolmesta modaalisuuden tyypistä, joihin kuuluvat episteeminen, deonttinen sekä dynaaminen modaalisuus.

Skottienglanti-nimitystä käytetään yläkäsitteenä kattamaan Skotlannissa esiintyvät kielen varieteetit skotista Skotlannin standardienglantiin. Koska sen sisältö on niinkin laaja, on sen tarkka määrittäminen monimutkaista. Skottienglannin modaalijärjestelmän on todettu eroavan melko suurestikin englanninenglannin vastaavasta, ja tämä tutkielma pyrkii osaltaan valaisemaan sitä, onko tilanne todellakin näin.

Teoria- ja metodiosuus tarkastellaan ensin tutkielman teoreettista viitekehystä, käyden läpi dialektologia ja korpuslingvistiikka. Tämän jälkeen esitellään tutkimusmateriaalina käytettävät korpuksat, *The Scottish Corpus of Texts and Speech* (SCOTS) sekä *The British National Corpus* (BNC). Kummastakin korpuksesta materiaalina on tutkielman aiheen huomioiden luonnollisesti puhutun kielen osa. Tutkimuksen rajaamiseksi korpuksista on käytetty systemaattista otantaa. Otoksen läpikäynnin ja analysoinnin jälkeen tulosten tilastollinen merkittävyys on vielä tutkittu χ^2 -testillä, jossa χ^2 on laskettu niin sanotun Pearsonin approksimaation avulla.

Tutkimuksen tuloksista selvisi että kolme verbeistä, *may, must* ja *ought to*, eivät esiinny dynaamisessa merkityksessä lainkaan kummassakaan varieteetissa. *Mayn* kohdalla kolme modaalisuuden tyyppiä esiintyvät samassa suhteessa sekä SCOTS:ssa että BNC:ssa, toisin kuin verbin *can* kohdalla, jossa verbin esiintymistiheyden suhde deonttisen ja dynaamisen modaalisuuden kesken oli päinvastainen SCOTS:ssa kuin BNC:ssa. Erot varieteettien välillä eivät kokonaisuudessaan tarkasteltuna näytä niin suurilta kuin saattoi olettaa, vaikka erot useimpien verbien kohdalla tilastollisesti merkittäviä ovatkin. Huomiota herättävin havainto oli kuitenkin tuplamodaalien täydellinen puuttuminen, sillä ne ovat yksi tärkeimmistä erottavista tekijöistä skottienglannin ja muiden englannin varieteettien välillä.

Avainsanat: Skottienglanti, korpuslingvistiikka, dialektologia, modaalisuus, puhekieli

Table of contents

1. Introduction	1
2. Scottish English then and now	3
2.1. The early stages	3
2.2. Inglis – Scottis - Scots	4
2.3. The present situation – Scottish English linguistic continuum.....	5
2.4. The problem with defining Scots – Dialect or language?.....	6
3. Data, theory and methods	9
3.1. Dialectology	9
3.2. Corpus linguistics	15
3.3. The corpora.....	18
3.3.1. The Scottish Corpus of Texts and Speech - SCOTS.....	18
3.3.2. The British National Corpus – The BNC.....	19
3.4. Methods used.....	20
4. Modality	23
4.1. Research on modal verbs.....	25
4.2. Modal verbs in Scottish English.....	27
4.3. Epistemic, deontic and dynamic meaning	31
4.4. Modal verbs	33
4.4.1. Can	33
4.4.2. Could.....	34
4.4.3. May	35
4.4.4. Might.....	36
4.4.5. Should	37
4.4.6. Must	38
4.5. Semi-modals	39
4.5.1. Ought to	39
4.5.2. Have to	40
5. Analysis	42
5.1. Can.....	42
5.2. Could	44
5.3. May.....	46
5.4. Might	49
5.5. Should.....	51
5.6. Must.....	53
5.7. Ought to	55
5.8. Have to	57
5.9. Conclusion.....	59
6. Bibliography	61

1. Introduction

(1) F1054 **Can** I just say a couple of wee things... (SCOTS)

The term of *Scottish English* as such refers to a broad range of regional and social varieties present in Scotland (Douglas 2006, 45). One area in which Scottish English is considered to differ from Standard English greatly, is the system of modal verbs (Miller 2004, 52).

In my MA thesis I will investigate the modal verbs *can*, *could*, *may*, *might*, *should* and *must* and the semi-modals *ought to* and *have to*. I examine how they are used in Scottish English compared to the usage of said modals in *English English*. My aim is to see in which of the three types of modality (*epistemic*, *deontic* and *dynamic*) the modals and semi-modals appear in both Englishes and whether there are any significant differences in their distribution. This examination will be conducted using two corpora – The British National Corpus (BNC) and The Scottish Corpus of Texts and Speech (SCOTS).

I will start by presenting the history of Scottish English, as it is important in understanding the current situation (Douglas 2006, 45). The issue of defining what Scottish English actually is, is far from clear – Fiona Douglas (2006, 45) states that Scottish English should be seen as a linguistic continuum with *Standard Scottish English* on one end and *Broad Scots* in the other. After presenting the history and current state of Scottish English and the difficulties in defining it, I will introduce the theoretical background, data and methods for my study. Chapter four will discuss modality taking an overall look on it, also introducing epistemic, deontic and dynamic modality. I will also present each chosen modal verb or semi-modal individually.

The research questions can be condensed as follows:

1. How are the modal verbs *can*, *could*, *may*, *might*, *should* and *must* and the semi-modals *ought to* and *have to* used in Scottish English as compared to English English?
2. Which of the three types of modality (epistemic, deontic and dynamic) do they represent and are there any significant differences in the distribution?

2. Scottish English then and now

Since the term *Scottish English* covers quite a range of varieties from *Broad Scots* to *Scottish Standard English* (Douglas 2006, 45), I find it important to explore the history and development of this variety of English so as to give a proper background for the situation as it is now. The aim of this section is thus to shed some light on how and why Scottish English became to be what it is today.

2.1. The early stages

The first appearance of a language derived from *Old English* in Scotland can be traced as far back as to 547, which marks the arrival of a group of Anglian invaders and the founders of the Kingdom of Bernicia which was situated around the same area that is today the Scottish-English border (Douglas 2006, 42). Upon their arrival the invaders also introduced their own language to this area where the language and culture had previously been Celtic (Douglas 2006, 42). Around the beginning of the mid-seventh century the kingdom had expanded to cover the area today known as the Scottish Lothians (ibid.).

Germanic language varieties of close relations to Old English came to the area in the eight century along with the Viking marauders who in the course of time ended up settling in Orkney and Shetland (Douglas 2006, 42). The influence of *Norn*, a Norwegian variety, is even today visible in *Insular Scots* (ibid.). Douglas (2006, 43) states that *Old Norse* has significantly influenced the English on both sides of the border, “evidenced by the adoption of Norse-influenced words”. She continues by noting that the influence was greater in Scotland and many words nowadays used in *Scots* have initially been loaned from Old Norse (2006, 43).

The movement of Anglo-Norman and Flemish overlords to Scotland was caused by the Norman Conquest in 1066 (Douglas 2006, 43). Along with them came also immigrant servants and retainers, which caused “a significant increase in the use of Anglo-Scandinavian throughout

lowland Scotland” (Douglas 2006, 43). According to Douglas (2006, 43) the “English” in Scotland was used mainly in the south and south-east of Scotland all the way to the twelfth century.

2.2. Inglis – Scottis – Scots

The Anglo-Scandinavian variety also known as *Inglis* spread its usage into an increasing amount of communicative functions and areas (Douglas 2006, 43). Instead of being confined to only spoken form, it began to be used in the written form as well (Douglas 2006, 43). She (2006, 43) also notes that Barbour’s epic poem *Brus* (1375) is the earliest document on the use of Inglis in written form. Due to the spreading of Inglis, “by 1390 Scottish Acts of Parliament began to be recorded in Inglis rather than Latin”, and by that time it had also become the dominant variety “for all Scotsmen to the south and east of the Highland line” (Douglas 2006, 43).

According to Douglas (2006, 43), *Scots* gained the position of the language to be used in formal registers between the 15th and 16th century. Douglas (2006, 43-44) also reminds us of the importance of remembering that the term Inglis first referred to the Anglo-Scandinavian varieties in use in Scotland and in England, and thus shows the close resemblance between these. The Scots did not differentiate their variety of Inglis until the late 15th century – the variety, to which the linguists refer to as *Older Scots*, was then named *Scottis* (Douglas 2006, 44). The period for *Older Scots/Scottis* is considered to reach from 1100 to 1700, followed by *Modern Scots* from 1700 onwards (ibid.).

Starting from the mid-16th century, Scots was beginning to be more and more threatened by the increasing Anglicization (Douglas 2006, 44). Examples of this were the introduction of an English Bible instead of the Scots one due to the Reformation in 1560 and the “shift towards English norms by many Scottish printers” in connection with the commencement of printing and the increased importation of English-printed books (Douglas 2006, 44). The status of Scots was further weakened by the Union of the Crowns, which brought about the move of the Scottish court to London thus causing many Scots writers to lose their patronage. Even though the

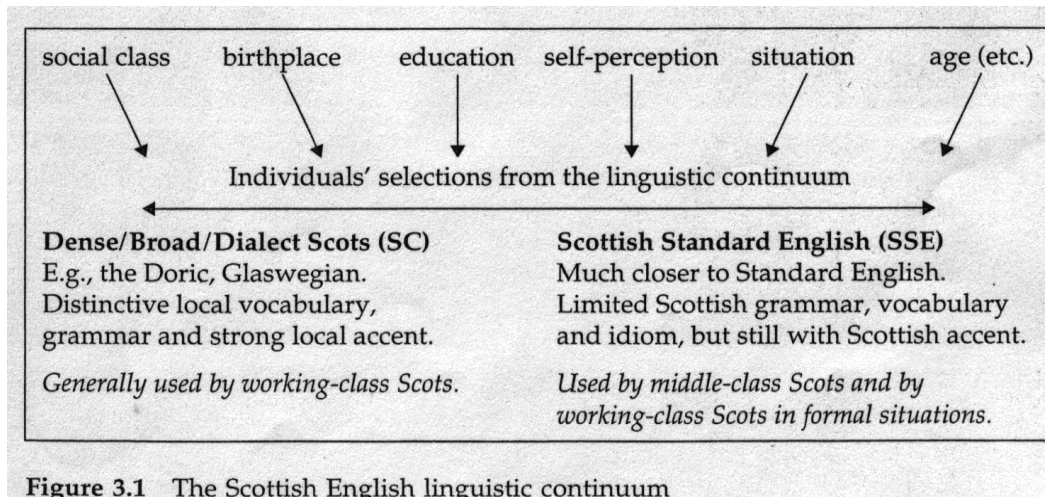
Anglicisation did damage in the written language, the spoken form of Scots remained clearly recognisable even for most of the 17th century (ibid.).

In addition to losing its social status at the Union of the Crowns and spiritual status at the Reformation, Scots was also deprived of its political status at the Treaty of Union in 1707 (Murison 1979, 9). Many members of the Scottish middle and upper classes also tried to remove Scotticisms from their speech and writing, and the speech of the middle class had acquired a significant amount of influence from southern Standard English, which later led to the development of Scottish Standard English (Douglas 2006, 44).

2.3. The present situation – Scottish English linguistic continuum

McClure (1994, 79) notes that *Scottish Standard English (SSE)* is originally a compromise between *Scots* and southern *Standard English*. This compromise is the result of native speaker's influence on a language they have learned and of the common conscious belief in the 18th century that the complete Anglicisation of one's speech is not preferable (McClure 1994, 79). Scottish Standard English currently enjoys the status of an autonomous and prestigious variety, which is now "recognised as an established national standard, throughout [sic] the English-speaking world" (ibid.). Scots on the other hand is considered to have low prestige, with the exception of literary contexts, and is mainly present in the speech of the working classes in Scotland (Douglas 2006, 45).

The aforementioned difference in the status of Scottish Standard English and Scots offers us the starting point for the idea of *Scottish English* as an umbrella term covering a linguistic continuum including regional and social varieties.



Picture 1. The Scottish English linguistic continuum. (Douglas 2006, 45)

At the “dense” end of the continuum is *Scots* (“localised forms known as ‘broad Scots’ and ‘dialect Scots’”) and the other end is represented by *Scottish Standard English* “a more or less homogeneous range of nationally acceptable norms of spelling, grammar, vocabulary and pronunciation” (McArthur 1979, 50). It is possible for individuals, influenced by outward factors such as social class, age, education etc., to move along the continuum, although it is noted that there is always the possibility that they are more drawn to one end than to the other (Douglas 2006, 45-46).

The *Scots* end of the continuum has more variation than the other end. *Scots* is often divided into the following groups: *Insular Scots*, *Northern Scots*, *North-East Scots*, *Mid Scots* (*East, West and South-West*), *Southern Scots* and *Ulster Scots* (Grant 1931, xlvii-xlviii). *SSE* in turn is used by people all over Scotland, with a slight coloration by the local features (Douglas 2006, 46).

2.4. The problem with defining Scots – Dialect or language?

As previously mentioned, the term *Scottish English* may be used to refer to all of the Scottish varieties, with Scots and Standard Scottish English acting as the opposite ends of the linguistic continuum consisting of the varieties (Douglas 2006, 45). According to Douglas (2009, 32-33), there have been attempts to standardise Scots, but they have not been particularly successful. Due to

this, Scots is still often left in the shadow of Scottish Standard English, which continues to enjoy its status as the prestigious form ordinarily used in formal situations (ibid.).

There is no universal agreement on whether Scots varieties are dialects of English or whether they are considered a separate language. When it comes to the view of Scots as a dialect, McArthur (2003, 139) points out that no matter how different Scots might be when it comes to vocabulary, syntax, spelling or sound it is still “not different enough structurally and lexically to be a language in its own right.” It is also noted that it does not have an official status nor is it considerably present in the teaching at schools. The role of Scots is considered minor or even nonexistent in Scotland’s legal system or administration (ibid.).

People who consider Scots to be a distinct language argue that it has “a highly distinctive sound system, grammar, and vocabulary, dating from the Anglian kingdom of Northumbria over a thousand years ago” (McArthur 2003, 139). Another point in defence of the view of Scots as a language is that it has “a varied and unbroken orthographic and literary tradition from the Middle Ages to the present day, including two medieval epic poems, ballads and love poetry” (McArthur 2003, 139). It has also been used to some extent by notable writers such as Robert Burns, Sir Walter Scott and Robert Louis Stevenson (McArthur 2003, 139). This view can also be supported by the notion that Scots has own dialects and “is recognized as a language by the European Bureau of Lesser Used Languages, an agency of the European Union” (ibid.). Due to this it is argued that as “equivalent forms in other countries have long been officially recognized as distinct languages, Scots should be so recognized and strengthened” (McArthur 2003, 140).

Even though there are many views on the certainty of the standing of Scots, there is still room for doubt. Some are not really certain what can be counted as Scots when it comes to its relationship to Standard English, and whether parts of it ought to be deemed “good” or “bad” (McArthur 2003, 140). This uncertainty may lead to unease, perhaps because of a “feeling that somewhere along the line they are losing access to many handy – even fine – traditional usages” (McArthur 2003, 140). Despite of this, they are usually aware of the “dialects, styles and writers” of

Scots (ibid.). Today it is a widely accepted view that while Scots may have once been a language on its own, it is now merely a dialect or a group of dialects (McArthur 2003, 141). For the purpose of this study, I will consider *Scots* as one of the varieties of *Scottish English* as I am examining Scottish English as a whole and Scots is part of it.

3. Data, theory and methods

In this section I will present the theoretical background for my study and introduce the data and methods used in analysing the data.

3.1. Dialectology

Dialectology can shortly be described as the study of dialects, both descriptive and theoretical (Newbrook 2002, 108). If the term '*dialect*' is interpreted in the broad sense to mean a 'variety of language', this field of study can be seen to be concerned with "analysing and describing related language varieties, particularly in respect of their salient differences and similarities" (ibid.).

Romaine (1994, 900) points out that dialectology is "a long-established branch of linguistics."

Dialectologists also develop theoretical frameworks for this kind of analysis and description to be able to form generalisations and hypotheses about the nature of linguistic differentiation and variation (Newbrook 2002, 108).

Newbrook (2002, 108) states that the form of dialectology known today began to form in the 19th century. This was preceded by a lengthy and widespread tradition known as *folk linguistics*. Folk linguistics was concerned with the "anecdotal and somewhat unsystematic discussion of regionalisms and variation in usage" (Newbrook 2002, 108). As a result of the continuation of this field, dialectology "has to deal with both theoretical and practical issues in respect of which folk linguistic concepts and beliefs have had, and continue to have, considerable currency" (ibid.). Due to this it is important to make a distinction between the views and definitions dialectologists have and those of lay commentators.

The main focus of dialectology in the 19th century was geographical, as linguistics was not socially oriented back then (Newbrook 2002, 110). Two major pioneering works in this field were published in the United Kingdom in 1889 by Ellis and in 1905 by Wright. The one by

Wright was also associated with the English Dialect Society, which was founded in 1873 (Newbrook 2002, 110). According to Romaine (1994, 900-901) the work of Georg Wenker is the starting point of traditional linguistic geography. Wenker composed 40 sentences with the idea to get information on differences in the German dialects (Romaine 1994, 900-901). This group of sentences was then sent to teachers who “were asked to transcribe the sentences using the speech characteristics of the area” (Romaine 1994, 901). This phase was followed by mapping the material and placing the different dialect features on the map (ibid.).

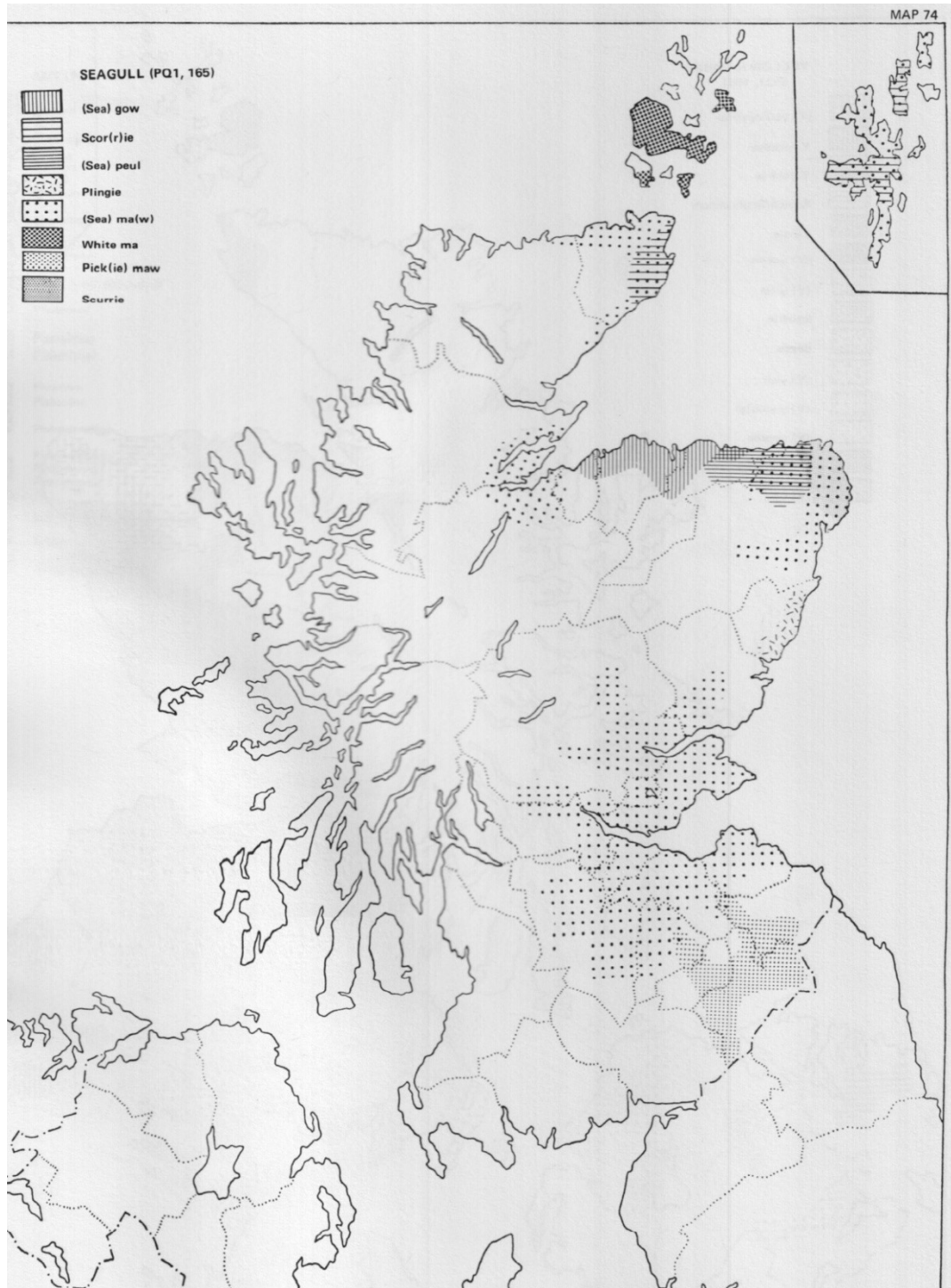
Another focus of interest in the early days of dialectology in its current form, now abandoned, was "the search for pure dialect, i.e. the supposedly regular and systematic form of speech produced by those remote from standardising influences" (Newbrook 2002, 110). The methodology used in this was the selection of *NORMS*, *Non-mobile, Old, Rural Males*, disregarding whether they actually represented the current usage of language in the community under research (Newbrook 2002, 110-111). The priorities in theoretical and methodological issues are quite different today which, together with a gradual shift of interest to syntax instead of phonology, lexis or morphology, has reduced the relevance of older publications (ibid.).

In connection of pioneering the concept of a *dialect atlas* in the 1870s, Georg Wenker and Ferdinand Wrede also developed frameworks for the methodology and analysis in fieldwork (Newbrook 2002, 111). Newbrook (2002, 111) notes that the main focus of their method was on indirect postal surveys, as opposed to the direct method involving face-to-face interviews used by Jules Gilliéron while surveying the French dialects in 1897. Gilliéron was inspired by the work Wenker had done and trained a fieldworker to collect material for him (Romaine 1994, 901). This material formed the French linguistic atlas, published in its entirety in 1910. The ongoing *Linguistic Survey of Scotland* in turn has used both of the aforementioned techniques (Newbrook 2002, 111). Another example of a linguistic atlas is *The Linguistic Atlas of the United States and Canada* (LAUSC), which was founded in 1930 (Chambers and Trudgill 1998, 17).

The notable *Survey of English Dialects* (SED) was started by Eugen Dieth and Harold Orton, with field research taking place from 1950 to 1961 leading to the publication of several volumes between 1962 and 1978 (Chambers and Trudgill 1998, 19). This survey was executed by the division of the country into four regions, these being the east and west midlands, the south and the north. The division was then followed by conducting interviews to gather material – the result being four volumes of *Basic Material* (Chambers and Trudgill 1998, 19). Interpretive volumes have also been published, with the aim of presenting the data on maps (ibid.). Some Finnish work on this area has also been done: early 1970s saw the launch of the regional English speech project at the University of Helsinki (Vasko, 2010). The eventual result of the project has been the *Helsinki Corpus of British English Dialects*, which was completed in 2006. Vasko (2010) states:

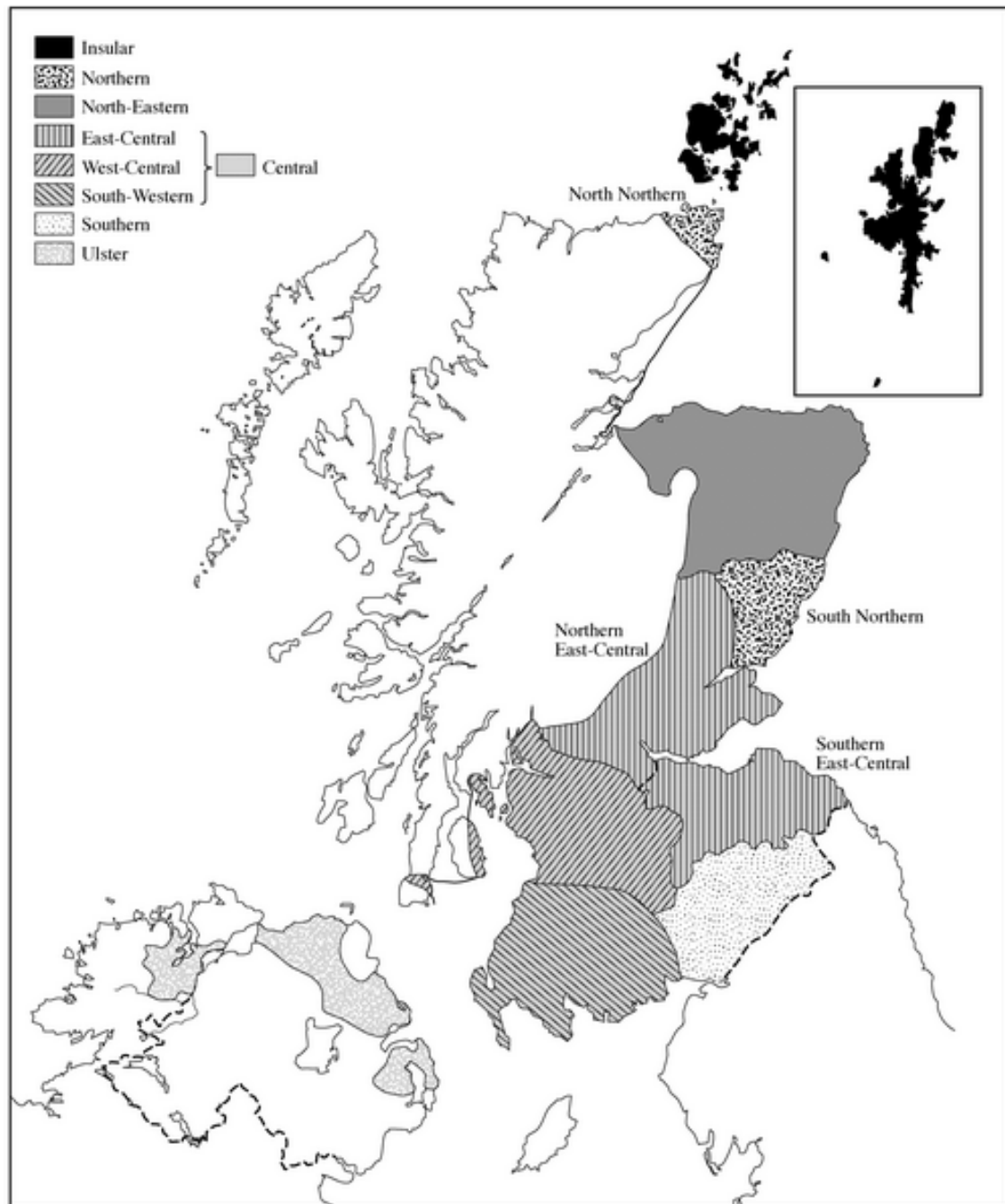
The aim of the Helsinki dialect project was to collect a corpus consisting of a substantial amount of continuous spontaneous speech to provide material for the study of dialect syntax, and thus to supplement the Leeds *Survey of English Dialects* (SED), which focuses mainly on phonological and lexical data.

A term closely linked to dialect atlases is *isogloss*, a line on a linguistic map separating two dialects from each other based on some linguistic feature (Yule 2006, 197). If the line formed by another linguistic item is found to follow the first one close enough, it is possible to draw another isogloss. When enough isoglosses go closely to each other on the map, they can be unified to form a clearer line which then indicates a *dialect boundary* (Yule 2006, 197). While isoglosses and dialect boundaries can give us a good view of regional dialects, they do not take into account the fact that dialects merge into one another instead of being sharply divided by a strict line (Yule 2006, 198). Nowadays the use of isoglosses has mainly been abandoned, and different symbols, which show the transition zones between two dialects a little more clearly and represent the linguistic reality better, are being used instead (Viereck 2005, 267). An example of the latter type of dialect map is the following map depicting the spreading of the word *seagull* and its synonyms:



Picture 2. Dialect map of the distribution of *seagull* and its synonyms in Scotland. (Mather and Speitel 1975, 165)

Another map provides us with the division of the Scottish dialects:



Map 10: The main dialect divisions of Modern Scots (based on CSD: Map 1; Gregg, 1985: Part II, Map 1).

Picture 3. The main dialect divisions of Modern Scots. (Macafee and Aitken, 2002)

Viereck (2005, 267) states that along with the comparative method used in the 19th century came two theories regarding the relationships between languages. The first of these was the *family tree theory*, a naturalistic concept by August Schleicher. The other, contrasting the family tree theory, was the *wave theory* by Johannes Schmidt (Viereck 2005, 267). The idea behind Schmidt's theory was that the differences increased when the distance between varieties increased. Current term for these situations is *geographical dialect continua*, which basically covers the idea that while dialects of a language that are very far away from each other geographically may not be mutually intelligible, they are still "linked by a chain of mutual intelligibility" (Viereck 2005, 267).

The early 20th century brought along the rise of structural linguistics, although it had fairly little influence on dialectology at first (Newbrook 2002, 111-112). Resulting from this the emphasis on synchronic systems came into regular use only after the 1950s. In the 1960s, a new tradition with the emphasis on obtaining more natural usage than that typical of questionnaire responses came into existence (Newbrook 2002, 111-112). This new tradition was pioneered in the United States by William Labov (ibid.).

1960s also introduced *generative dialectology*, with its main emphasis on providing formal descriptions of variation "within some form of the generativist paradigm" (Newbrook 2002, 113). 'Interactionist' models and analyses of linguistic variability, "in which mere correlation of linguistic and non-linguistic phenomena has a more marginal role" and which concentrated on the active role of the participants in a conversation with the aim of constructing the identity of the participants through the (culturally significant) choices made from the range of variants at their disposal, were developed in the 1980s and 1990s (Newbrook 2002, 113).

Newbrook (2002, 113) states that the more recent dialectological research has been characterised by the continuously expanding use of computers. The new urban dialectology emphasises openly spoken, modern varieties instead of "obsolescent and obscure forms of speech" (Newbrook 2002, 113). The theoretical orientation has also strayed away from the very descriptive

approach to dialect studies present in the beginning (ibid.).

3.2. Corpus linguistics

The methodological framework for this thesis is formed by the field known as *corpus linguistics*. I will first discuss *corpus linguistics* as a whole and then move onto the little more specific definition of *comparative corpus linguistics*, also taking into account the aspect of spoken corpora, as spoken language is the subject of my study.

The simplest way to describe a *corpus* is to consider it “a body of texts” (Anderson and Corbett 2009, 4). In corpus linguistics there is, however, often the need to be more precise and a corpus is described as “a large principled collection of texts, that is, one which has been created for a purpose” (ibid.). Leech et al. (2009, 24) shortly define corpus linguistics as “the study or analysis of language through the use of (computer) corpora”. According to Anderson and Corbett (2009, 9), the history of corpus linguistics is long, if the field of study is understood to be “simply the empirical analysis of language” - texts have been analysed manually for research purposes for over a century.

According to Meyer (2002, 1), *the Brown corpus*, created in the 1960s, was the first computer corpus around. This field of study can be seen to have existed since then despite the fact that the term *corpus linguistics* itself “was apparently not in use until the 1980s” (Leech et al. 2009, 24). Other important landmarks in the development of corpus linguistics have been the creation of the *LOB Corpus* (Lancaster-Oslo/Bergen Corpus) and the work done on the *Survey of English Usage*, which is to be thanked for “major grammatical descriptions of English” (Anderson and Corbett 2009, 9).

Leech et al. (2009, 24) call the creation of the Brown corpus “the landmark event for the development of corpus linguistics in the modern sense”. McCarthy and O’Keeffe (2009, 1012) have noted that according to various views on how a corpus ought to be designed, a corpus should be “a principled collection of texts that is assembled for a specific purpose”. A modern corpus

should be a representation of “naturally occurring language, in electronic form” with the design of representing a language as a whole or parts of it (Anderson and Corbett 2009, 4).

McCarthy and O’Keeffe (2009, 1008) state that the earliest spoken corpora were in many cases developed to research child language acquisition. According to them (2009, 1009), many spoken corpora have also been created to accompany written corpora, as is also the case with The British National Corpus (BNC) and The Scottish Corpus of Texts and Speech (SCOTS) that are used in this study – they include both written and spoken texts. In the beginning many of the spoken corpora were in English, but nowadays there exists spoken corpora for many other languages as well, such as French, Greek and Egyptian Arabic, to mention a few (McCarthy and O’Keeffe 2009, 1010).

If we compare written and spoken corpora, the spoken ones have a tendency to be smaller than their written equivalents due to problems with transcription and collection, and the lack of time and funds (McCarthy and O’Keeffe 2009, 1012). Another difference is that whereas with written corpora the collection of texts is pretty straightforward, in compiling spoken corpora the creators have to face the dilemma of deciding what actually can be constituted as a text – spoken texts do not have similar clear orthographic boundaries as the written ones (ibid.).

Anderson and Corbett (2009, 99) note the following:

Until the widespread availability of automatically searchable, digitised, language corpora, linguists had to rely for their observations on more limited language data, manually collected and analysed, or alternatively, they had to rely on intuition, their reflections on their own knowledge of language and their feelings about what is acceptable and unacceptable, and what particular constructions mean.

Corpus linguists are of the opinion that the study of language on the kind of scale enabled by electronic corpora, brings out structures and behaviour that would not be so easily studied based on intuition solely (Anderson and Corbett 2009, 99). Despite this it is still a fact that intuition plays its part in the analysis, in the selection of the features to be searched or in constructing hypotheses, and when actually analysing the data (Anderson and Corbett 2009, 100).

Comparative corpus linguistics, or as Schmied (2009, 1141) calls it, *contrastive corpus studies*, essentially means comparing two languages with the help of corpora. This definition can also be extended to cover different varieties of the same language (Schmied 2009, 1141). The history of comparing languages goes way back in history, and despite being one of the important influences in the development of philology in the 19th century, was to a large extent neglected in the beginning of the 20th century (Schmied 2009, 1142). It should be noted, however, that “the unification of corpus methodology and contrastive linguistics has not led to a new sub-discipline of contrastive corpus linguistics... but rather to a new wave of corpus-based contrastive studies” (Schmied 2009, 1142).

The corpus has to be a fair sample of the language it represents – otherwise the possibility of making sound generalisations based on the data can be compromised (Anderson and Corbett 2009, 5). In addition to this, it is important to remember that “a corpus can only provide positive evidence of a usage or construction” (Anderson and Corbett 2009, 5). Forms that are never used will not be present in a corpus (*ibid.*).

Meyer (2002, 11) notes that corpora can be very good resources for various types of research. This is exemplified by the mention that creating dictionaries is more effective due to studies on large linguistic corpora. Corpora have also shown the way to “new areas of research” and brought “new insights to traditional research questions” (Meyer 2002, 11). In Anderson and Corbett’s (2009, 22) view the “surfeit of easily accessible, seemingly transparent evidence is one of the great virtues of corpus linguistics.” Another strength in “corpus-informed language study” is the possibility to see in detail the construction of conversational English (Anderson and Corbett 2009, 91-92). Corpora provide us with the possibility to construct larger pictures on language exceptions and patterns with much more detail (Anderson and Corbett 2009, 122). Corpus evidence can also be said to reduce “the researcher’s bias and prejudices about language” (*ibid.*).

3.3. The corpora

The two corpora chosen for this study – *The Scottish Corpus of Texts and Speech* (SCOTS) and *The British National Corpus* (BNC) – were chosen due to their representativeness of those varieties of English I wanted to examine. The BNC offers a wide range of material (Anderson and Corbett 2009, 10-11) which made it an obvious choice for the representation of English English in my study. SCOTS in turn is to my knowledge the only wider corpus to concentrate solely on Scottish English, which was the reason for choosing it. I will now present both the corpora individually and in more detail.

3.3.1. The Scottish Corpus of Texts and Speech – SCOTS

The SCOTS Project is the first project made in a large scale of its type in Scotland (SCOTS).

The Scottish Corpus of Texts and Speech (SCOTS) was first introduced online in November 2004, and by May 2007 it had reached a total of four million words of text. Currently the corpus consists of 1300 spoken and written texts, 77% of which is written and 23% spoken texts “in the form of an orthographic transcription, synchronised with the source audio or video” (SCOTS). The creators of the corpus have aimed at including whole texts whenever possible, although this has not always been possible due to copyright issues (SCOTS). The corpus offers “extensive sociolinguistic metadata” in connection of all of the texts, which can be used when filtering the search (Anderson and Corbett 2009, 15). Its Advanced Search feature allows the browsing of the documents and downloading them in bulk to the users own computer as plain text files, which can then be analysed with search software such as WordSmith Tools (Anderson and Corbett 2009, 15).

The aim of this corpus is to cover texts from 1945 to the present day, latest documents dating from 2011 (SCOTS). Despite this aim, most of the content is from “the latter part of this period”, with the majority of the spoken texts having been recorded in the 21st century and

“recorded specifically for the SCOTS project” (SCOTS). The corpus has tried to do justice to the wide range of varieties between Scots and Scottish English, with material “from as wide a range of geographical locations as possible” and with speakers and writers of all kinds of backgrounds (SCOTS). Anderson and Corbett (2009, 14) note that SCOTS “includes texts in Scottish English and varieties of contemporary Scots, plus a few texts in Scottish Gaelic.” The aim of the SCOTS corpus can be said to be “to represent language as it is actually used”, although despite this aim it is not “a truly representative corpus” due to difficulties of obtaining certain types of material, and permission and copyright issues, as stated above (SCOTS). The corpus is not currently grammatically annotated (*ibid.*). I will access the corpus directly from their website.

3.3.2. The British National Corpus – The BNC

The *BNC* consists of 100 million words of both written and spoken language collected from a wide variety of sources with the aim of representing a large section of British English from the latter half of the 20th century, as also stated by McCarthy and O’Keeffe (2009, 1009) who noted that the BNC was designed to “represent as wide a range of modern British English as possible”. This means that it does also include Scottish English – a point which influenced the choice of spoken language as the subject of examination. In the spoken part of the BNC it is possible to restrict the search results to include or leave out speakers of a certain variety of English if needed when using the BNCweb interface – this is not possible in the written part. When doing the searches I restricted the results to exclude Scottish English.

Anderson and Corbett (2009, 10) describe the BNC as the “gold standard among corpora of British English” due to “its large size, level of annotation and availability”. Both the written and the spoken part are further divided into smaller, more precise categories or genres (BNC). Some examples of the type of material included in the written part are extracts from regional and national newspapers, academic books, popular fiction, and essays. The spoken part, which takes up 10 % of the data in the corpus, includes spoken language collected in different

contexts, and transcripts of unscripted informal conversations, recorded by volunteers from different classes (BNC). One of the important features of the BNC is that it is “automatically part-of-speech-tagged” (Anderson and Corbett 2009, 11).

The building of the BNC began in 1991 and continued until 1994 (BNC). “The first general release of the corpus for European researchers” took place in 1995 (BNC). This was followed by a phase of improving the tagging of the corpus, a process led by Geoffrey Leech, Roger Garside and Tony McEnery. Before the publication of the second edition, the corpus also went through “correction and validation of the bibliographic and contextual information in all the *BNC Headers*” (BNC). This edition, known as *BNC World*, was made available to the public in 2001. On the BNC website it is stated that “additional mark-up for lemma and simplified word-class annotation was added” along with the improvement of the way multi-word units are treated (BNC). The improved version, known as *BNC XML Edition*, was published in 2007 (ibid.). The version used in this study is version 4.3, published in 2010. It is accessed through BNCWeb.

3.4. Methods used

The selection of the verbs was based on a formation of pairs – *can* and *could*, *may* and *might*, *should* and *ought to* and *must* and *have to*. All of these pairs are connected to each other in that their meanings are approximately synonymous.

After having selected the two, previously presented, corpora and deciding which verbs to take into scrutiny, I started the collection of the data. The first step was to decide the size of the sample – I decided on 100 occurrences per verb which would then make a total of 1600 units to examine. The selection of the occurrences from the data was done using *systematic sampling* as the sampling design. Lavrakas (2008, 871) notes that systematic sampling “is a random method of sampling that applies a constant interval to choosing a sample of elements from the sampling frame.” According to Luojola (2006, 25) the first unit is chosen using a random number, after which the rest of the units are picked using a constant interval. I calculated the random number using the

calculator at Random.org, and got the result of 18 – the first unit then being the 18th one. The following units were picked using the interval of five until I reached 100 units, with the exception of the verbs *might* and *should* in SCOTS, where the interval was four in order to enable the accumulation of 100 units as the interval of five would have lead to an insufficient sample.

The sampling was followed by going through the units and marking each unit as epistemic, deontic or dynamic. The statistical significance of the results was tested using the *chi-square* (χ^2) test. The chi-square test can be used as a test of independence or as a goodness-of-fit test (Lavrakas 2008, 95). In the case of testing the independence it ”measures the significance of the relationship between two categorical variables, representing the first step toward bivariate analysis” (Lavrakas 2008, 95). According to Lavrakas (2008, 95) the logic is ”to calculate the distance between the observed frequencies within the contingency table and the condition of statistical independence.” The calculation is done using the so-called *Pearson's chi-squared test*, the formula for which is as follows:

$$h = \sum_{i=1}^k \frac{(f_i - np_i)^2}{np_i}$$

(Ruohonen 2011, 40)

In the formula f_i represents the amount of occurrences of one of the modality types in SCOTS and np_i represents the corresponding amount of occurrences in BNC. k is the number of degrees of freedom in the h , in other words the number of different modality types with actual occurrences in each verb. The calculation is applied to each of the three types of modality per verb and the resulting chi-square values are then added together to get the chi-square for the total amount. After the chi-square value is calculated, it is compared to the table value with the help of k . The table can give a rough estimate of the value of p . According to Nuzzo (2014, 151) a p value ”measures whether an observed result can be attributed to chance.” This is then continued by noting that the result is statistically significant if the value is 0,05 or smaller, as conventionally considered (ibid.). The exact p value can be calculated using computer software – in this case the calculations were

performed using Microsoft Office Excel 2007 and the accuracy of these was then verified by performing the same calculations in an online calculator¹ and comparing the results.

¹ <http://www.socscistatistics.com/tests/chisquare/>

4. Modality

In this section I will take a look at *modality*, first discussing the subject in general. This is followed by a section on research on the area. Then I will present the use of modals in the syntax of Scottish English and the different types of modality, proceeding after that to present the modal verbs and the semi-modals chosen for scrutiny in this thesis.

Carter and McCarthy (2006, 638) define modality as “a speaker’s or a writer’s attitude towards, or point of view about, a state of the world.” The range of meanings that can be expressed with the modal verbs is wide and consists of ideas such as permission, obligation, necessity and ability (Biber et al. 1999, 73). Modals are used when discussing the truthfulness of something: whether it is true or real, or whether it should be considered mere speculation instead of being taken as certain knowledge (Carter and McCarthy 2006, 638). Politeness and formality are also often expressed by using modal items (ibid.)

Perhaps the clearest way to express modality is by the use of *modal verbs* (Carter and McCarthy 2006, 638). According to Carter and McCarthy (2006, 638) there are nine so-called *core modals* in total: *can, could, may, might, will, shall, would, should* and *must*. These verbs are referred to as *central modal auxiliary verbs* by Biber et al. (1999, 483). As an exception to the list offered by Carter and McCarthy (2006, 638) and Biber et al. (1999, 483), Swan (2005, 325) also adds *ought* to the set.

In addition to core modals, there is a group of verbs known as the *semi-modals* (Carter and McCarthy 2006, 638). They act similarly to the actual modal verbs but have some properties in common with lexical verbs (Carter and McCarthy 2006, 638). Carter and McCarthy (2006, 638) list the following as belonging to semi-modals: *dare, need, ought to* and *used to*. Leech et al. (2009, 92) in turn add forms like *have to* and *be able to* into this category, despite the fact that they are not considered to be actual semi-modals by for example Carter and McCarthy (2006, 663; 670) who list *have to* under other modal verb phrases and *be able to* in modal expressions built with the verb *be*.

Biber et al. (1999, 73) in turn consider semi-modals as a combination of the marginal auxiliaries *dare*, *need*, *used to* and *ought to* and the multi-word verbs close to modal verbs in their meaning, such as *be supposed to*, *(had) better*, *have to*, *be going to* and *(have) got to*. The difficulty in defining which of the forms are actual semi-modals and which are not is probably best described by the following quotation: “The word 'semi-modals' is not a precise term. It refers to a loose constellation of verb constructions...” (Leech et al. 2009, 91).

Modal verbs differ remarkably from other verbs in that they only have one form and do not take inflections for number or person (Carter and McCarthy 2006, 639). This means that they do not, for example, have the third person singular -s, as exemplified by “She **may** know his address. (NOT ~~She mays...~~)” (Swan 2005, 325). They do not have infinitives nor do they have *-ing* form or *-ed* participle form thus being deficient in progressive and perfect forms (Carter and McCarthy 2006, 639). Their place in the verb phrase is in the beginning of the phrase, followed by another verb, such as an auxiliary verb, lexical verb or the substitute verb *do*, in its base form:

- (2) We **might stay** an extra night.
- (3) We **should be leaving** soon.
- (4) It **might have got** lost in the post. (Carter and McCarthy 2006, 639-640)

The following verb cannot, however, usually be another modal verb (Carter and McCarthy 2006, 639), although this is actually possible in some regional varieties (Biber et al. 1999, 483), a good example of this being Scottish English (Miller 2004, 53). Swan (2005, 325) notes that *do* is not used when forming tags, short answers, negatives or questions with modals.

Leech et al. (2009, 91) call semi-modals “probably the most cited cases of grammaticalization in the ongoing history of English.” Alternative names for semi-modals are *quasi-modals*, *periphrastic modals* (Biber et al. 1999, 484), *non-typical* or *marginal modal verbs* (Carter and McCarthy 2006, 657). Whereas the core modals cannot be marked for tense or person, this is possible for the semi-modals (Biber et al. 1999, 484). They can also occur as non-finite and it is

possible for the infinitive forms of semi-modals to occur together with another semi-modal or with one of the core modals (Biber et al. 1999, 484).

4.1. Research on modal verbs

In their study, Leech et al. (2009, 71) have examined the changes in the frequency of modal verbs in written English. The research was done using several corpora: *the Brown University Corpus* (Brown), *the Freiburg-Brown Corpus* (Frown), *the Lancaster-Oslo/Bergen Corpus* (LOB) and *the Freiburg-Lancaster-Oslo/Bergen Corpus* (F-LOB). In addition to using these corpora, the British English spoken subcorpora from *the Diachronic Corpus of Present-Day Spoken English* (DCPSE) was also utilised, as were the British National Corpus (BNC) and *the Longman Corpus of Spoken American English* (LCSAE) (Leech et al. 2009, 71).

As the corpora were of approximately the same size, raw counts were used instead of normalised frequency counts (Leech et al. 2009, 72). The modal verbs under inspection were the core modals, with *ought to* and *need(n't)* added as “peripheral members” and the comparison was done using the data from 1961 and 1991/2 (Leech et al. 2009, 72). The results showed that there was a clear decline of frequency, the overall percentage being 10,6% - for individual verbs it varied from 2,2% to 43,5% (Leech et al. 2009, 73). The only verb to break the pattern was *can*, with a rise of 1,3%. They also note that the “order of frequency of the modals is the same in both 1961 and 1991/2...” (ibid.).

Leech et al. (2009, 73-75) then took this further and examined whether the decline varied between different text types and found that of four subcorpora, Learned, General Prose, Fiction and Press, there had been no decline at all in the Learned subcorpus. Instead the frequency of modal verbs had risen in this subcorpus. In Leech et al.’s (2009, 76-77) comparison of spoken English to written English, they found that the frequency order of the modal verbs varied a little between spoken and written language. The frequency of modals was also noted to be clearly higher in the spoken data. Their conclusions bring to notion the overall decline of frequency but also “a

tendency toward monosemy: toward the increasing prevalence of one meaning over others” (Leech et al. 2009, 89-90). It is also pointed out that the decline appears to be greater in the case of marginal usages (ibid.).

As opposed to the results on modal verbs, Leech et al. (2009, 98) found in their similar, accompanying study on semi-modals, that their frequency had experienced noteworthy increase from 1961 to 1991/2. They do, however, point out that despite this, modal verbs are still “more than five times as frequent as semi-modals” (ibid.).

In her MA thesis, Rajalahti (2006, 2) studies Philippine English and Singapore English “through modal and quasi-modal verbs, or more precisely, those of obligation and necessity.” The verbs chosen for the study are modal verbs *should*, *must* and *need* and quasi-modals *ought to*, *had better*, *need to*, *be supposed to*, *have to* and *have got to* (Rajalahti 2006, 4). The study is corpus-based, and the five corpora used are *The International Corpus of English* (ICE) in the variety corpora of Philippine English, Singapore English and British English, *The Santa Barbara Corpus of Spoken American English* (SBC) and the Freiburg-Brown Corpus (Frown) (Rajalahti 2006, 5).

Rajalahti (2006, 8-9) uses the division into root and epistemic modality, with the emphasis on the root meaning (obligation and logical necessity). In ICE, two categories were chosen for the analysis: S1A (private dialogue) and W2F (creative writing), these two being “the best to exemplify the possible independent developments in PhiE and SinE compared to AmE and BrE” (Rajalahti 2006, 14). In the case of the SBC, the analysis was done on Part I; in the Frown the categories “K (general fiction), L (mystery and detective fiction), M (science fiction), N (adventure and western) and P (romance and love story)” were chosen (Rajalahti 2006, 24-25).

Rajalahti (2006, 78) found that most of the verbs appeared more frequently in Philippine English and Singapore English than in American English or British English, the only exceptions being *have got to* and *ought to*, which were more frequent in British English than in Singapore English. The most frequent type of subject with the verbs was personal pronouns – in

Singapore English the subject was implied instead of actually stating it in the sentence (Rajalahti 2006, 79). When expressing obligation, the most common verbs used were noted to be *should*, *have to* and *must*, with *need*, *ought to*, *had better*, *have got to*, *supposed to* and *need to* on the decline. All of the verbs appeared surprisingly little in the Frown (ibid.). Rajalahti (2006, 81) states that of the text categories that were studied, “private dialogue, instructional writing and persuasive writing were the most prolific in using the modals.”

The findings proved that root modality was significantly more common in all varieties than epistemic modality – the only exception to this was the verb *must* (Rajalahti 2006, 82). The reason suggested for this is that due to the very hierarchical social system, “it may be that it is more acceptable to oblige people in the Philippines and Singapore” (Rajalahti 2006, 82-83).

4.2. Modal verbs in Scottish English

Jim Miller (2004, 52) notes that “the system of modal verbs in Scots is massively different from that of Standard English.” These differences were noted as early as in the 18th century by writers like Hume, Sinclair, Mitchell and Beattie, and modal verbs were high on their lists of the so called Scotticisms (Beal 1997, 365). All of them give the idea of Scots forms being the “wrong” or incorrect ones, more or less explicitly stated in their Scotticism lists (Beal 1997, 365-366). As Miller refers to Scots or Broad Scots, this mainly applies to a very small part of the linguistic continuum under the term *Scottish English*. However, it is probably safe to assume that some of these rules will be present in other forms on the continuum as well, if not necessarily in *Scottish Standard English* which represents the opposite end of the continuum from *Scots*.

The first distinctive feature of the system of modal verbs in Scots is the absence of the verbs *shall*, *may* and *ought* (Miller 2004, 52). According to Miller (2004, 52) there are no occurrences of any of these verbs in *ECOSSE* (the Edinburgh Corpus of Spoken Scottish English), although they do appear in formal announcements and in writing from time to time, exemplified by the following notice: “This shop **shall** be open on Monday” (ibid.). Beattie remarks that “the Scots

are more apt to misapply *will* than *shall*, especially in the first person singular and plural” (Beattie, quoted in Beal 1997, 366). Beal (1997, 366) suggests that this remark is probably due to “the absence of *shall* and *should* from Scots, as it is only in the first person that *shall/should* are 'unmarked', according to the 'traditional' rules for these verbs.” However, Beal (1997, 366) does note that since other dialects have “caught up” with Scots when it comes to this feature, it cannot necessarily be said to be distinctively Scottish anymore. There is no notion of *shall* being absent in Grant and Dixon (1921, 117-118) as they note *shall* being found in the forms of *sal* and *sall* and *sal* also being shortened to *'se* or *'s*.

When expressing permission, *may* is replaced with *can*, *get to* or *get + gerund*: “You **can** have this afternoon off” (Miller 1993, 116-117). Grant and Dixon (1921, 120) consider *may* as normally being the equivalent of *can*. Beal's (1997, 367) notions undermine the status of *can* used to express permission as a Scotticism as well, as *may* is no longer that much present even in standard English due to being “increasingly confined to more formal usage in standard English”. The forms presented for *can* in Grant and Dixon (1921, 120-121) include, in addition to *can canna* (“cannot”), *cud* (“could”) and *couldna*, *cudna* or *cwidna* for “could not”. It is also stated that *can* and *could* “are used after the auxiliaries 'will' and 'have' in place of 'be able', 'been able': but not in the Northern dialects” (Grant and Dixon 1921, 121).

Should replaces *ought* in Scots, although *want* is also frequent in this context: “You **want** to come out and attack right away” (Miller 1993, 117). The meaning of *must* is restricted in Scots in comparison to its usage in Standard English (Miller 2004, 52). When in Standard English it is used to express both conclusion and obligation, in Scots it is reserved only for denoting conclusion: “You **must** be exhausted” (Miller 2004, 52). *Must* in its negative form *mustn't* is used in the sense of ‘I conclude that not’, as in “This **mustn't** be the place” (Miller 2004, 53). The same meaning is expressed by *can't* in Standard English. In marking obligation in Scots its place is taken by *have to* or *need to* (ibid.). Miller (1993, 117) adds that there is also often a distinction made between *must* and *have to* in describing Standard English: *must* expresses internal or self-

compulsion, while *have to* denotes external compulsion. This distinction, however, is not relevant when discussing the use of these verbs in Scots (ibid.). Grant and Dixon (1921, 121) present the Scottish modal verb *maun*, which is used in the place of *must*, and its negative forms *maunna*, *mauna* and *manna*, replacing *must not*.

According to Miller (1993, 118) the verb *need* “behaves like a main verb in Scots.”

This can be exemplified by the following sentences:

(5) You **needn’t** leave immediately. (Standard English)

(6) You don’t **need** to leave immediately. (Scottish English)

In addition to acting like a main verb, *need* is also used in the progressive, as in “They’re **needing** to paint the window” (Miller 1993, 118). Grant and Dixon (1921, 123) note *need* as having past tense and past participle form *not*. *Need* expresses obligation, and is used when indicating external compulsion, just like its equivalent *have to* (Miller 1993, 118).

In addition to the already mentioned verbs for expressing obligation, it can also be indicated by *supposed to* and *meant to* (Miller 1993, 119). Furthermore, *meant to* has the meaning ‘it is said that’, exemplified by “The new player is **meant to** be real fast” (Miller 1993, 119). The verb *will* denotes future tense in spoken Scots:

(7) You **will** have the money tomorrow. (Miller 2004, 52)

Apart from this, it appears in interrogatives (Miller 2004, 52). The notion of *will* denoting future tense is supported by Grant and Dixon (1921, 117) who state that “‘will’ is the ordinary auxiliary form interrogative for the future tense; ‘shall I’, ‘shall you’ are not used.” Grant and Dixon (1921, 116-117) note the forms for *will* as being *wull* (“will”), *winna*, *wonna* (“will not”), *wad*, *wud* (“would”), *wadna*, *widna* and *wudna* (“would not”).

Perhaps the most notable difference in the use of modal verbs between Standard English and Scots is the occurrence of double modals in Scots:

- (8) He'll **can** help us the morn/tomorrow.
- (9) They **might could** be working in the shop.
- (10) She **might can** get away early. (Miller 1993, 119)

In addition to being combined with *can* or *could*, as in the examples, *might* also frequently occurs with *would* and *should*, although still not as frequently as with *can* or *could* (Miller 1993, 120). It has been suggested that *might* is developing into an adverb, being “syntactically equivalent to *maybe*” (ibid.). Double modals have in fact “become 'distinctively Scottish' by virtue of their disappearance from Southumbrian dialects” (Beal 1997, 335). A relatively old double modal sequence is *will can*, which has first been mentioned in 1915 by Wilson and in 1921 by Grant and Main-Dixon, although it appears to be in decline (Miller 1993, 120; Miller 2004, 54). This tendency was visible in the results of an Honours dissertation by McIver, where it was found that while people over 60 in Orkney used this construction, those under the age of 25 “neither used it nor recognised all the combinations” (Miller 2004, 54).

As opposed to Standard English, modal verbs can appear after the infinitive marker *to* in Scots: “You have **to can** drive a car to get that job” (Miller 1993, 120; Miller 2004, 53-54). Beal (1997, 368) states that Standard English has considered double modals to be ungrammatical since the Early Modern period, owing to “radical changes in the English auxiliary verb system which effectively outlawed any non-finite forms of modal verbs”. Beal (1997, 368) finds it strange that 18th century lists of Scotticisms have no notions of double modals despite the fact that the construction is bound to have seemed odd to speakers of Standard English. The earliest mention of double modals in The Scottish National Dictionary (SND) is from the 19th century (ibid.). It is possible that the apparent absence of any earlier evidence of double modals is just a coincidence as they are deemed to have been a firm part of spoken Scots “by the beginning of the nineteenth century for Scott to employ them”. (Beal 1997, 368-369) Beal (1997, 369) goes on to note that as

the double modals and non-finite forms of *can* have existed “at least from the late 18th century”, it is a sign that the Scots modal system is indeed different from the Standard English one.

4.3. Epistemic, deontic and dynamic meaning

Modal meaning can be divided into several types. Collins (2008, 130) presents the “tripartite distinction” made by Palmer, which divides the meanings in three – *epistemic*, *deontic* and *dynamic*. Palmer's division is also the one I will be using in this study. There are also other ways to group the modal meanings, a well-known one being the division into *root* and *epistemic modality* made by Coates (Collins 2008, 130). Aarts et al. (2014, 366) describe *root modality* simply as consisting of “any kind of modality that is not epistemic modality.” The modal type of *alethic modality* is also presented by Aarts et al. (2014, 20) and it is stated to usually go under epistemic modality. In the case that the distinction between the two is made alethic modality deals with logical deduction, whereas epistemic modality is concerned with confident inference (ibid.). Modality can also be divided into three types by using the groups alethic, epistemic and deontic (Aarts et al. 2014, 114).

Epistemic modality has to do with “the truth of the proposition” (Collins 2008, 130). It concerns certainty, probability and possibility – likely facts are evaluated and this is then followed by making conclusions or predictions based on these evaluations (Carter and McCarthy 2006, 638). Biber et al. (1999, 485) note this type of modality as referring to “the logical status of events or states, usually relating to assessments of likelihood: possibility, necessity, or prediction”. Palmer (1979, 41) considers this type of modality as

the simplest to deal with. In both its syntax and its semantics, it is the kind of modality that is most clearly distinct from the others and has the greatest degree of internal regularity and completeness.

According to Palmer (2003, 7) “epistemic modality is concerned solely with the speaker's attitude to status of the proposition”. Some scholars also use the term *extrinsic* to refer to this kind of modality (Biber et al. 1999, 485).

Palmer (1986, 51) notes that the term *epistemic*, which applies to modal systems indicating the notions of possibility and necessity, should instead apply “to any modal system that indicates the degree of commitment by the speaker to what he says” and should especially take into account evidence of the type of hearsay and report. Palmer (1986, 51) is aware that this usage might be “wider than usual” but considers it etymologically justified as the origin of the word *epistemic* derives “from the Greek word meaning 'understanding' or 'knowledge'” and would thus allow the interpretation of “showing the status of the speaker's understanding or knowledge.”

The second type, *deontic modality*, is also known by the name of *intrinsic modality* (Biber et al. 1999, 485). This type of modal meaning has to do with controlling the course of events and “getting things done” (Carter and McCarthy 2006, 639). In addition to expressing obligation, this type of modality conveys whether something is considered permitted, forbidden, necessary or desirable (ibid.). Deontic modality is considered to concern actions and events controlled directly by humans (Biber et al. 1999, 485). According to Collins (2008, 130), deontic modality is “concerned with conditions upon an action deriving typically from an external source.” This view is supported by Palmer (2003, 7), who considers deontic modality to be directive, meaning that the control of the event is not in the grip of the subject in the sentence. Palmer (1979, 58) states that it is not as easy to distinguish deontic modality from the other types as it is to make the difference between epistemic modality and the others.

Dynamic modality is connected to an individual's ability or volition to do something (Collins 2008, 130). In bipartite divisions of modal meaning it is generally combined with either epistemic or deontic modality, as exemplified by the well-known division into root and epistemic meanings by Coates (Collins 2008, 130). Huddleston and Pullum (2002, 178) note that “the clearest cases of dynamic modality are concerned with properties and dispositions of persons.” As opposed to deontic modality, in dynamic modality the control can be seen to be internal to the subject (Palmer 2003, 7). Palmer (1979, 36-37) further divides dynamic modality into *neutral dynamic modality* and *subject oriented modality*. Of these, neutral dynamic modality is used to indicate

things that are “possible for” or “necessary for”, whereas subject oriented modality concerns the ability and the willingness of the subject (ibid.).

4.4. Modal verbs

In this section I will present the modal verbs I have chosen for this study as they are used in Standard English. The six chosen verbs are *can*, *could*, *may*, *might*, *should* and *must*.

4.4.1. Can

Can is the most frequent of the modal verbs and is most often used to express permission and ability (Carter and McCarthy 2006, 642). In addition to being used to give, seek or deny a permission and to indicate what someone is capable of, *can* may also be used to express general truths, in other words “statements about events and states which are true or which are usually the case (ibid.):

- (11) **Can** you speak French? (ability)
- (12) You **can** stop work early today. (permission)
- (13) **Can** I help you? (offer)
- (14) I **can** read Italian, but I **can't** speak it.
- (15) These roses **can** grow anywhere.
- (16) Scotland **can** be very warm in September. (Swan 2005, 97-98).
- (17) Steel **can** resist very high temperatures. (Carter and McCarthy 2006, 642).

A fourth possible use for *can* is the indication or evaluation of “logical possibilities”, such as in the following sentences:

- (18) That **can't** be right.
- (19) How **can** they be here already? They only left ten minutes ago? (Carter and McCarthy, 643).

Swan (2005, 98) remarks that *can* may be used to express what is possible or impossible for us to do owing to the surrounding circumstances:

(20) We **can** go to Paris this weekend, because I don't have to work.

(21) There are three possibilities: we **can** go to the police, we **can** talk to a lawyer, or we **can** forget all about it.

Can may also be used when discussing future actions possible due to present abilities, circumstances or decisions (Swan 2005, 99):

(22) I've bought the tent, so we **can** go camping next weekend if we want to.

(23) She **can** win the race tomorrow if she really tries.

4.4.2. Could

Could is most frequently used for expressing that something is possible or probable and in suggestions (Carter and McCarthy 2006, 643):

(24) I think that **could** be the answer to the problem.

(25) We **could** all be having holidays on the moon within thirty years.

(26) I **could** just cook dinner tonight and then you **could** cook it some other night and, you know, reciprocate. (Carter and McCarthy 2006, 643).

(27) It **could** rain this afternoon.

(28) When you've finished the washing-up you can clean the kitchen. Then you **could** iron the clothes, if you like. (Swan 2005, 97-101).

Another regular use for *could* is in requests and orders or in asking for a permission (Swan 2005, 101):

(29) **Could** you lend me five pounds until tomorrow?

(30) Do you think you **could** help me for a few minutes?

(31) **Could** I ask you something? (Swan 2005, 101).

(32) **Could** I talk to you for a moment?

(33) **Could** you give me a ring if you can't make it, Bob. (Carter and McCarthy 2006, 643-646).

In addition to asking for a permission, *could* can also be used when discussing "permission that has already been given or refused" as in the sentence "She said I **could** come as often as I liked" (Swan

2005, 101). *Could* is considered to be more polite and formal in this use than *can*. The two verbs are also connected in that *could* may be used as the past form of *can*, or as a more tentative alternative to *can* (Swan 2005, 97; 101):

- (34) When I was younger I **could** play tennis very well. (Swan 2005, 97).
- (35) When I was a kid I **couldn't** swim at all. I only learnt when I was thirty.
- (36) We **could** hear that dog barking all through the night. (Carter and McCarthy 2006, 644).

Combined with *have* + past participle form, *could* can be used to express ability or opportunities in the past that did not turn into reality. The same structure is also used when criticising people (Swan 2005, 99):

- (37) I **could have married** anybody I wanted to.
- (38) I **could have won** the race if I hadn't fallen.
- (39) You **could have helped** me – why did you just sit and watch? (Swan 2005, 99).

Another structure for criticising or expressing disapproval is *how could you*, as in “**How could you** forget that we're going out to dinner tonight?” (Carter and McCarthy 2006, 644).

4.4.3. May

The modal verb *may* can be used to grant, deny or ask for a permission much like *can*. *May* can also refer to weak probabilities, that something is not very likely to happen (Carter and McCarthy 2006, 644):

- (40) **May** I see that?
- (41) There's a bank holiday in between, so it **may** or **may not** get to you by the end of that week. (Carter and McCarthy 2006, 644).
- (42) **May** I put the TV on?
- (43) Students **may not** use the staff car park. (Swan 2005, 318).

While *can* generally refers to something the speaker believes to be true, or at least to be the case most of the time, when expressing probability, *may* is used to refer to possible events. In relation to general truths, *may* is used to refer to things that are likely to happen or normally do, and can as such be seen as the “formal equivalent of *can*” in this connection (Carter and McCarthy, 645):

(44) Frog spawn **may** be found in river beds at that time of year. (Carter and McCarthy 2006, 645).

(45) We **may** go climbing in the Alps next summer.

(46) I think Labour are going to win. ~ You **may** be right. (Swan 2005, 316).

Other use for *may* is the usage in formal expressions of wishing something good to someone and in formal curses (Carter and McCarthy 2006, 645):

(47) **May** you both have a long and happy married life together.

(48) **May** you rot in hell! (Carter and McCarthy 2006, 645).

(49) **May** the New Year bring you all your heart desires. (Swan 2005, 319).

May can also have a concessive meaning, and this is especially the case if accompanied by *well* or *but*: “I **may** be in danger of stating the obvious, **but** I shall state it anyway” (Carter and McCarthy 2006, 645).

4.4.4. Might

According to Carter and McCarthy (2006, 646) *might* could be described as the more “indirect and tentative” version of *may*. Swan (2005, 316) describes *might* as the “the less definite or more hesitant form of *may*, suggesting a smaller chance – it is used when people think something is possible but not very likely”:

(50) Joe **might** come with me. (Swan 2005, 316).

Most often *might* is used to express probability and possibility of something happening:

(51) And I **might** buy a video camera as well. (Carter and McCarthy 2006, 646).

(52) Peter **might** phone. If he does, ask him to ring later.

(53) Where's Emma? ~ I don't know. She **might** be out shopping, I suppose. (Swan 2005, 316).

Might may also be used to express permission – it is then considered rather formal and polite and mainly used in indirect structures (Swan 2005, 318):

(54) I wonder if I **might** have a little more cheese. (Swan 2005, 318).

(55) **Might** I speak to Mrs Lutterworth? (Carter and McCarthy 2006, 647).

Although *might* is not often used as the past form for *may*, it does replace *may* in indirect speech:

(56) I **may** go to Scotland. ~What? ~I said I **might** go to Scotland. (Swan 2005, 317).

The structure *might as well* is used, as noted by Swan (2005, 319), to “suggest that one should do something because there is nothing better, nothing more interesting or nothing more useful to do.”

The same structure is also used with *may*:

(57) There's nobody interesting to talk to. We **may as well** go home.

(58) Shall we go and see Fred? ~OK, **might as well**. (Swan 2005, 319).

4.4.5. Should

Carter and McCarthy (2006, 653) state as the most frequent uses for *should* the indication of what is considered to be “the ideal or desired state of affairs”:

(59) He **should** have been here at five and he's not here yet.

(60) She **should** be wearing glasses.

(61) I think the authorities **should** do it and there **should** be, you know, guidelines from a central place. (Carter and McCarthy 2006, 653).

(62) Everybody **should** wear car seat belts.

(63) Applications **should** be sent before December 30th. (Swan 2005, 510).

Should is also used to express that something is possible or probable, logically or due to it being normal in the circumstances (Swan 2005, 510):

- (64) She's away, but she **should** be back tomorrow.
- (65) Henry **should** get here soon – he left home at six.
- (66) We're spending the winter in Florida. ~That **should** be nice. (Swan 2005, 510).

When an event is surprising or hard to believe, *should* is used (Carter and McCarthy 2006, 653):

- (67) I'm sorry that he **should** be so upset by what I said.
- (68) I'm amazed that he **should** have done something so stupid. (Carter and McCarthy 2006, 653).
- (69) It's surprising that she **should** say that to you.
- (70) I was shocked that she **shouldn't** have invited Phyllis.
- (71) I'm sorry you **should** think I did it on purpose. (Swan 2005, 512).

Should can appear with *if* in conditional clauses in formal contexts (Carter and McCarthy 2006, 654). The same type of meaning can be expressed with *should* + *in case* or *so that* (Swan 2005, 510):

- (72) **If** you **should** see Caroline, tell her I've got the tickets.
- (73) I'll get a chicken out of the freezer **in case** Aunt Mary **should** come.
- (74) He turned the radio down **so that** he **shouldn't** disturb the old lady. (Swan 2005, 510.)

4.4.6. Must

Must is often used when expressing “different strengths of obligation”, whether it be a polite invitation or a strict order (Carter and McCarthy 2006, 654):

- (75) You **must** get those done by tomorrow.
- (76) I told her she **must** keep her door locked at all times.
- (77) All passengers **must** present valid photo identification at check-in for all flights. (Carter and McCarthy 2006, 654-655).
- (78) Plants **must** get enough light and water if they are to grow properly.
- (79) You **must** be here before eight o'clock. (Swan 2005, 335).

Swan (2005, 335) adds to this the emphatic invitations, such as "You really **must** come and see us soon."

Other frequent use for *must* is in connection of deduction or to express reproach (Carter and McCarthy 2006, 654-655):

- (80) I'm twenty-eight, so she **must** be twenty-seven.
- (81) **Must** you have that music so loud? (Carter and McCarthy 2006, 654-655).
- (82) I'm in love. ~ You **must** be very happy.
- (83) There's the doorbell. It **must** be Roger.
- (84) You **must** be Anna's sister – you look just like her. (Swan 2005, 334).

Carter and McCarthy (2006, 655) note that *must* does not have a past form, so in expressing past time it has to be replaced by *had to*. It is also often recommendable to use *have to* instead of *must* when referring to future obligations (ibid.):

- (85) If he turns up after midnight, then he'll **have to** eat whatever he can find in the fridge. (Carter and McCarthy 2006, 655.)
- (86) When you leave school you'll **have to** find a job.
- (87) I **had to** cycle three miles to school when I was a child. (Swan 2005, 335-337).

When expressing the negative form of *must* in deducing something, *must* is replaced by *can't* or *cannot* (Carter and McCarthy 2006, 656):

- (88) That **can't** be right. No it isn't. (Carter and McCarthy 2006, 656).
- (89) She's not answering the phone. She **can't** be at home. (Swan 2005, 334).

4.5. Semi-modals

The selected semi-modals, *ought to* and *have to* are presented here.

4.5.1. Ought to

Meaningwise, *ought to* is similar to *should* (Carter and McCarthy 2006, 659). It is therefore used when expressing for example states of affairs that are considered ideal or desired:

(90) I really **ought to** go outside and get some fresh air for a bit.

(91) You **ought to** put more money in your pension fund. (Carter and McCarthy 2006, 659).

According to Carter and McCarthy (2006, 660) *ought to* can be used to say what is probable or likely to happen. It can also be used in interrogatives, but it is very rare in that context and mainly used in formal styles:

(92) I think it **ought to** take about three hours, if the traffic is not too bad.

(93) Er, look at Brinton and if necessary go to Rowland's book. But I think Brinton **ought to** be able to give you the information.

(94) Who do you think it is? **Ought we to** call the police? (Carter and McCarthy 2006, 660).

Ought to does not have a past form and Carter and McCarthy (2006, 660) state that the “perfect construction *ought to have* + *-ed* participle is used to refer back to states of affairs which were desirable at points in the past”:

(95) We probably **ought to have talked** about it ages ago.

(96) They **ought to have told** you, didn't they? (Carter and McCarthy 2006, 660).

4.5.2. Have to

Have (got) to is similar to *must* in its meanings, being used to refer to “obligation and deductions” (Carter and McCarthy 2006, 667). When expressing obligation, *have (got) to* is generally used when the obligation is external to the speaker (ibid.):

(97) I **have to** be in at six every morning.

(98) I'll be back in a minute. I've just **got to** make a phone call.

(99) We've **got to** stay over a Saturday night to get the cheap flight.

(100) **Have you got to** get up early tomorrow? (Carter and McCarthy 2006, 667).

(101) Do you often **have to** travel on business?

(102) Sorry, I've **got to** go now. (Swan 2005, 210).

Swan (2005, 210) points out that *have (got) to* can be used when talking about a future obligation that is already fixed, otherwise the future form would be *will have to*:

(103) I've **got to** get up early tomorrow – we're going to Devon.

(104) One day everybody **will have to** ask permission to buy a car. (Swan 2005, 210).

(105) I've **got to** go to the dentist at half past ten tomorrow.

(106) I've chipped a bit off one of my teeth. I'll **have to** go to the dentist. (Carter and McCarthy 2006, 667).

Like *must*, *have (got) to* can be used in making assumptions or in drawing conclusions (Carter and McCarthy 2006, 668):

(107) He said a white house next to the village shop. Er, this **has got to** be it.

(108) So I said, 'Maureen, this one'll be our train,' and of course it **had to** be the wrong one. So we got off at the next stop. (Carter and McCarthy 2006, 668).

5. Analysis

5.1. Can

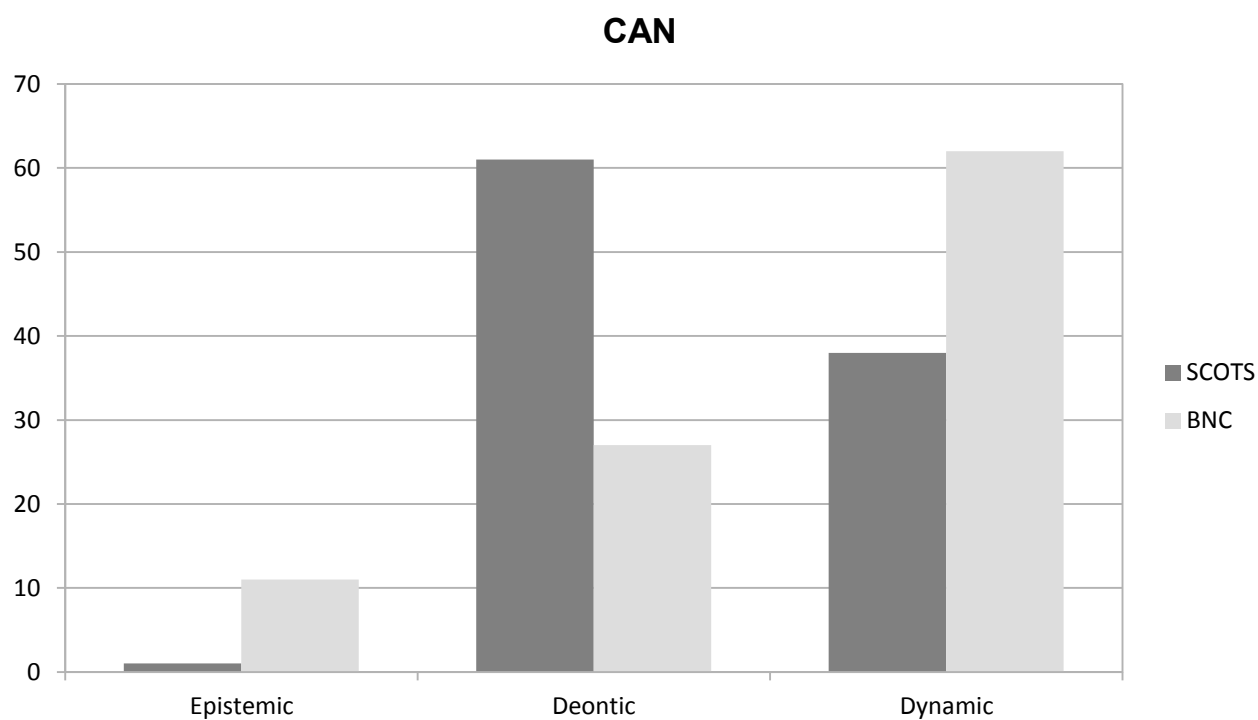


Figure 1. Modal verb *can* in Scottish English (SCOTS) and English English (BNC).

Can				
	SCOTS	BNC	Total	Chi-square
Epistemic	1	11	12	9,09
Deontic	61	27	88	42,81
Dynamic	38	62	100	9,29
Total	100	100	200	61,20
$p=5,15 \times 10^{-14} < 0,05$				

Table 1. Modal verb *can* in Scottish English (SCOTS) and English English (BNC).

As previously stated, Carter and McCarthy (2006, 642) note that the modal verb *can* is most often used to express permission and ability. This notion is supported by the findings on the part of English English, as dynamic modality which concerns ability and volition is the most frequent type of modality present in the results from the BNC occurring 62 times out of 100. However, the amount of hits representing the type of modality that concerns permission, deontic modality, is surprisingly low in light of the notion and this type is represented by only 27 units out of 100:

- (109) KCE 3366 **Can** you un-- o-- can you undo it? [unclear] turn it? DYNAMIC
- (110) KB8 2001 **Can** you see all the sand on the grass? DYNAMIC
- (111) KBW 13386 You **can** both have a go [pause] DEONTIC
- (112) KBW 10730 Yes, you **can** have half. DEONTIC (BNC)

In the case of Scottish English these two groups are the other way round, with deontic *can* being the most common in the hits from the SCOTS corpus by 61 units out of 100 and followed by 38 units per 100 of dynamic *can*:

- (113) F1054 **Can** I just say a couple of wee things... DEONTIC
- (114) F1122 //A rectangle// **Can** I do the shapes now? DEONTIC
- (115) F1121 Right, hold on a second. **Can** you manage off that chair? DYNAMIC
- (116) M1098 **Can't** find mair cars. Cars. DYNAMIC (SCOTS)

The overall profile of the findings per corpora is quite different, although a common feature for both of them is the low amount of epistemic *can*, BNC having 11 units and SCOTS only one.

Statistically the difference is still significant, with the p value being $5,15 \times 10^{-14}$, which is smaller than 0,05. Dynamic *can* is clearly more common in English English than in Scottish English (BNC 62, SCOTS 38). With deontic modality the situation turns around, with deontic *can* being clearly more common in Scottish English than in English English (BNC 27, SCOTS 61).

5.2. Could

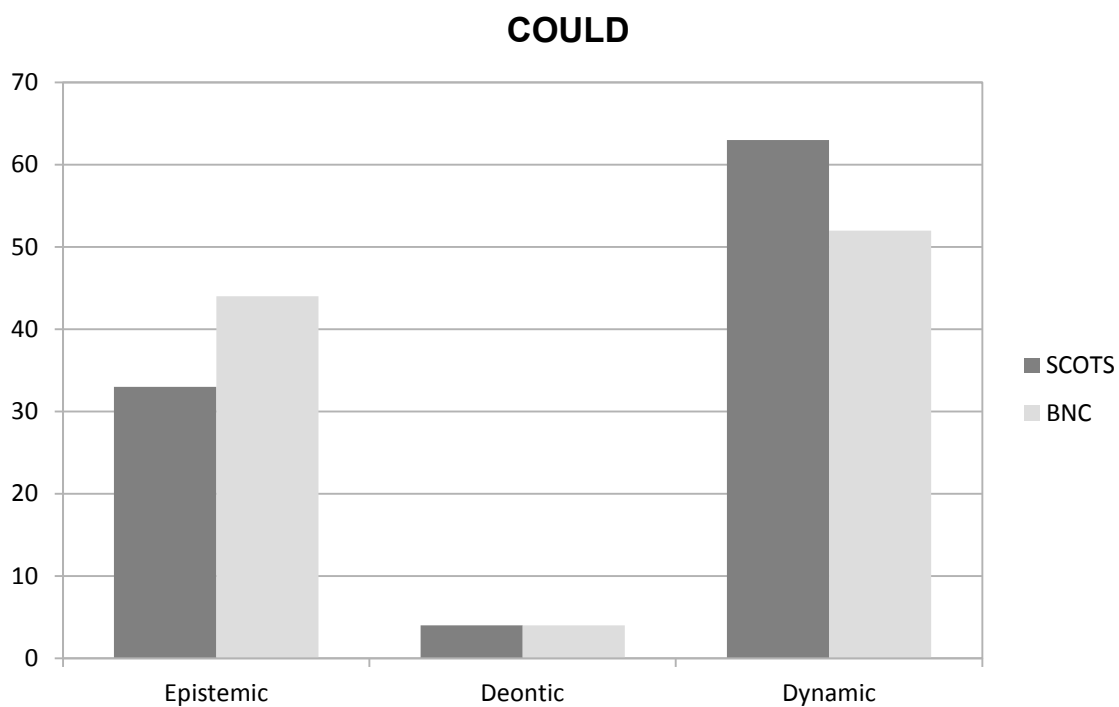


Figure 2. Modal verb *could* in Scottish English (SCOTS) and English English (BNC).

Could				
	SCOTS	BNC	Total	Chi-square
Epistemic	33	44	77	2,75
Deontic	4	4	8	0
Dynamic	63	52	115	2,33
Total	100	100	200	5,08
p=0,08 > 0,05				

Table 2. Modal verb *could* in Scottish English (SCOTS) and English English (BNC).

Following Carter and McCarthy (2006, 643), *could* expresses most often possibility and probability. This notion is supported by the results to the extent that the occurrences of epistemic *could* are quite frequent in both corpora (33 of 100 in SCOTS and 44 of 100 in BNC):

(117) KD8 6642 Yeah and the lad from work's a scouser he's said it, I'm going home now to watch the football, I think he's going there actually, I think he **could** be there, Rob you know him. EPISTEMIC

(118) KCN 3538 Well you **could** take them back on Friday. EPISTEMIC

(119) HYK 713 Well, I can't con-- I, I think that **could** be very attractive, myself with the trellis work and roses [pause] EPISTEMIC

(120) KD2 491 I don't believe they **could** get lost coming round here not with all the back fields on fire. EPISTEMIC (BNC)

(121) F1192 **Could**'ve been. Ehm and it's got a name like, when you hear, when you hear the name it sounds like it would be all Catholic, it's, what's it called? On the cover it's got a, a photo of a roof. EPISTEMIC

(122) M1015 port an the chances are they **could** eh thingmy, but I have tae say eh I've been at sea for what forty forty-odd years now an I can, it's only the the the last generation, the last ten years that I've seen earrings. EPISTEMIC (SCOTS)

However, epistemic modality is not the type most frequently represented in the data – this honour goes to dynamic modality, with 63 occurrences of 100 in SCOTS and 52 of 100 in BNC:

(123) KBW 4069 I know we're not inviting anybody for lunch Christmas day [pause] **couldn't** cope with that but it would be quite to have them [pause] [unclear] in the afternoon [pause] [unclear] can't cope with the people! DYNAMIC

(124) KCU 4548 Cos little Jane **couldn't** get as clo-- , if I sing they might DYNAMIC

(125) KE1 671 He **couldn't** remember a single [unclear] ! DYNAMIC (BNC)

(126) F1043 Abb wool it was cried. It was like a- an off-white but y- you **could** actually feel the oil fan //fan you were knittin wi it your hands were// DYNAMIC

(127) M1010 ke-, 'kerihaundit', eh that came frae the Kerrs at Fernyhurst eh they were supposed to be eh left-handit. And they had the tower an they had this this this stair in the in the opposite direction so if they were attacked by onybody away back in the Reivers times they **could** fight left-haundit on their stair cause it was shapit a different way. DYNAMIC

(128) F1026 but it wisnae too bad in that he had an aviary, so you **could** watch all these //bonnie birds, [laugh]// DYNAMIC (SCOTS)

As there are no notions of *could* being used differently in Scottish English as compared to English English, apart from being noted to appear in double modals together with *might* (Miller 1993, 119-120), there is no point in trying to explain any possible differences with Scotticisms. In fact, the distribution of the occurrences between the three types of modality in the corpora is quite similar. Dynamic modality is the most frequent, followed by epistemic modality. Both samples had only four occurrences of deontic *could*. The findings in both corpora were very similar throughout the sample, and the p value being bigger than 0,05 (0,08) shows that the difference is not statistically significant.

5.3. May

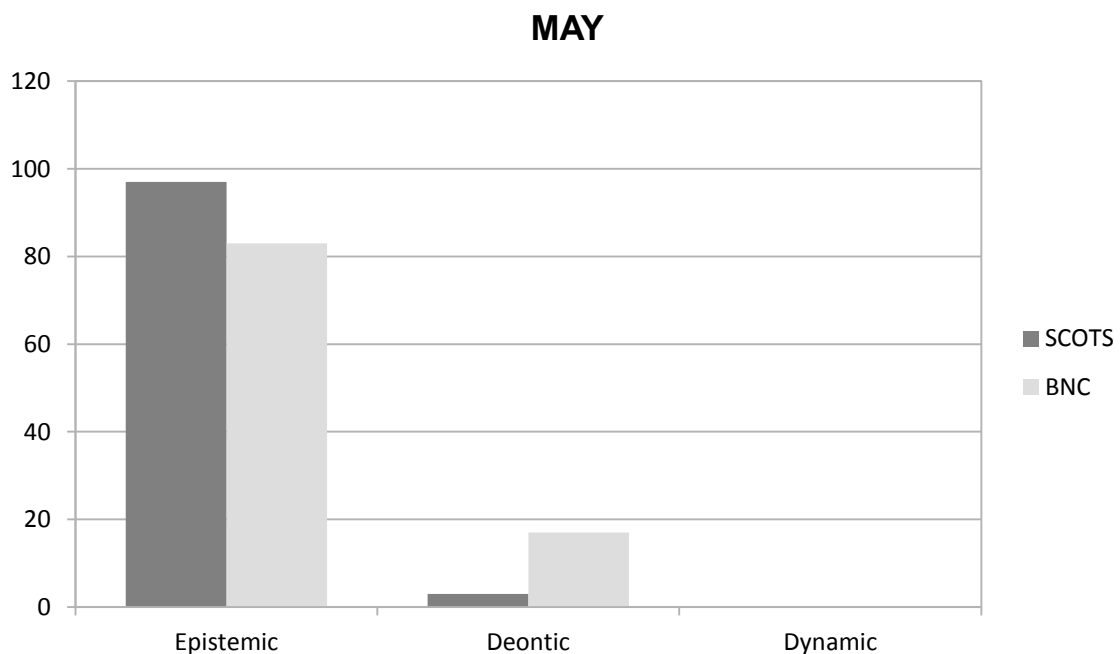


Figure 3. Modal verb *may* in Scottish English (SCOTS) and English English (BNC).

May				
	SCOTS	BNC	Total	Chi-square
Epistemic	97	83	180	2,36
Deontic	3	17	20	11,53
Dynamic	0	0	0	-
Total	100	100	200	13,89
p=0,00019 < 0,05				

Table 3. Modal verb *may* in Scottish English (SCOTS) and English English (BNC).

May is generally used to express granting, denying or asking permission (deontic modality) or the probability of something (epistemic modality) (Carter and McCarthy 2006, 644-645). Both of these types are represented in the results, unlike dynamic modality which does not occur at all – due to this there is naturally no chi-square for dynamic modality.

There are not particularly many occurrences of deontic *may* - 3 units out of 100 in SCOTS and 17 out of 100 in BNC:

- (129) HFO 123 But I think it's a little bit rich if I **may** say so... DEONTIC
- (130) KCK 788 No you **may** not draw a picture. DEONTIC (BNC)
- (131) F640 ...to be excused from the table, "please **may** I// DEONTIC
- (132) M087 ...and so I'd like to end eh, if I **may**, eh by by reading you eh... DEONTIC (SCOTS)

Epistemic *may* is clearly the most common type of *may* in both corpora, with 97 occurrences in SCOTS and 83 occurrences in BNC:

- (133) KLV 948 Obviously if business goes really swimmingly then it **may** be easier to [unclear], I'm not sure. EPISTEMIC
- (134) KGW 100 It's also the case that, as I was saying earlier, people **may** be erm may be may be attributing. EPISTEMIC (BNC)

(135) F1680 //But I-, you want the honest truth, and this **may** sound terrible.

EPISTEMIC

(136) F1189 //Mmhm.// Do you think she **may** have changed her mind, or did she
//settle there?// EPISTEMIC (SCOTS)

In addition to expressing permission and probability, *may* can be used in formal expressions that are used to wishing something good to someone (Carter and McCarthy 2006, 645). This usage was also exemplified in the findings:

(137) F639 ...to throw it in the fire, and say “Lang **may** yer lum reek”,
“Long may your chimney //smoke”,// EPISTEMIC (SCOTS)

In the case of Scottish English there are two things that ought to be noted concerning the findings. Miller (2004, 52) has stated that *may* is absent in Scots except for some rare occurrences in writing and formal announcements. This statement seems to be proven wrong by the findings which show that *may* is very much present in spoken Scottish English. Miller (1993, 116-117) has also pointed out that *may* is replaced with other verbs when expressing permission. This notion is supported by the relatively few occurrences of deontic *may* in the data. The differences in the distribution of the types of modality between Scottish English and English English are proven to be statistically significant with a p value of 0,00019.

5.4. Might

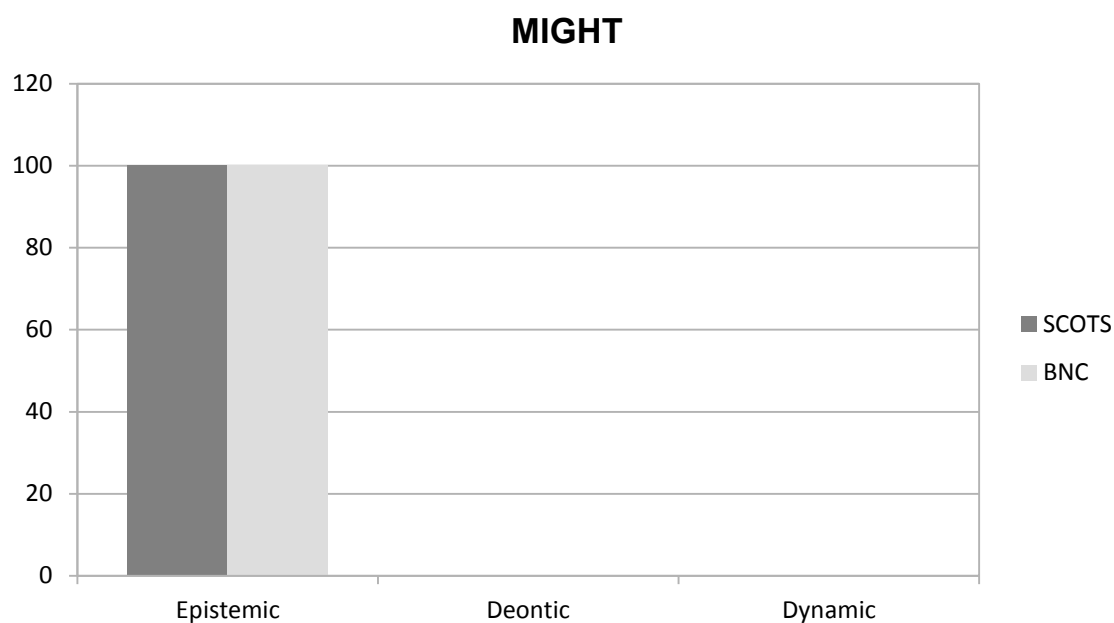


Figure 4. Modal verb *might* in Scottish English (SCOTS) and English English (BNC).

Might				
	SCOTS	BNC	Total	Chi-square
Epistemic	100	100	200	-
Deontic	0	0	0	-
Dynamic	0	0	0	-
Total	100	100	200	0
p=1 > 0,05				

Table 4. Modal verb *might* in Scottish English (SCOTS) and English English (BNC).

The only type of modality present in the findings is epistemic modality – the sample of 100 units from both corpora did not include any other type of modality:

(138) KBF 11991 So they **might**, they might work us. EPISTEMIC

(139) KD8 4306 Cos you've got to allow for like they **might** not have done it before.
EPISTEMIC

(140) KE3 5872 What **might** be an idea is to get her some good blotters. EPISTEMIC (BNC)

(141) M1012 an it depends how how ye say it ehm yer inflection could change the word 'gey', cause ehm if ye ask somebody if they enjoyed it, they might say it was 'gey guid' or they **might** say 'gey guid', an that would change the meanin, first yin would be, 'gey guid' would be eh enjoyed it // 'gey guid' would mean 'it was alright'! // EPISTEMIC

(142) F1009 Work box to me sounds a bit more feminine perhaps than than a kit of tools that a // a builder **might** use. // EPISTEMIC (SCOTS)

This is in line with Carter and McCarthy's (2006, 646) notion that *might* is most often used in expressing probability and possibility. According to Miller (2004, 53) *might* could be developing into an adverb with a meaning equivalent to *maybe*. This notion also supports the clear domination of epistemic modality.

Might could also be used to express permission, although this use is considered very formal (Swan 2005, 318). The formality of *might* in expressing permission could be the explanation behind the total lack of deontic *might* in the findings, as spoken language is likely to be less formal than written language. The distribution is the exact same in both corpora, in which case the difference is not statistically significant.

5.5. Should

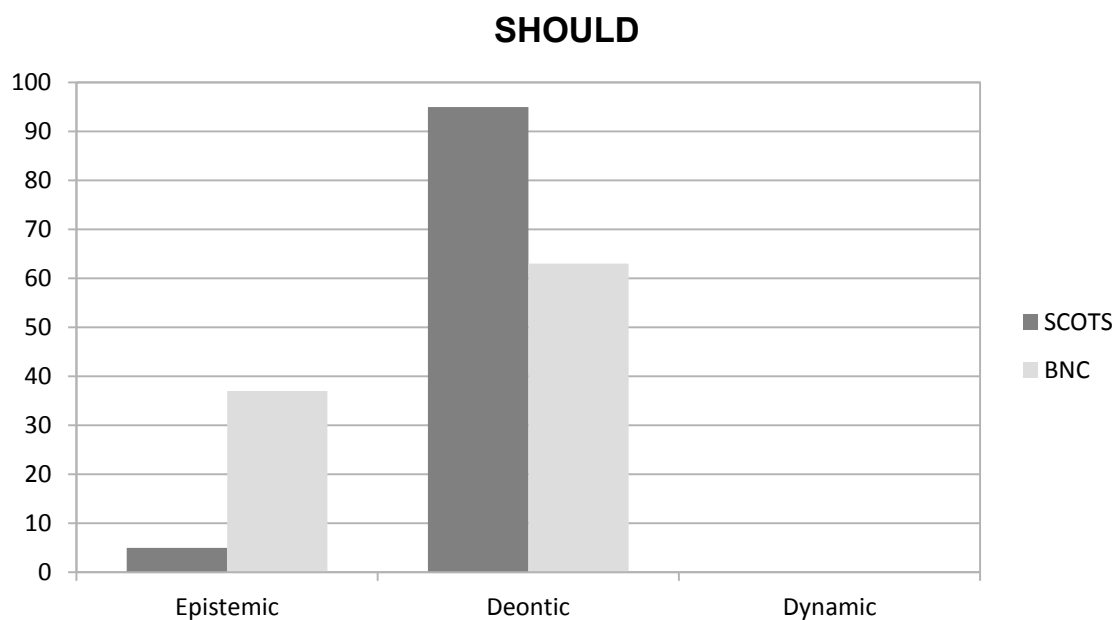


Figure 5. Modal verb *should* in Scottish English (SCOTS) and English English (BNC).

Should				
	SCOTS	BNC	Total	Chi-square
Epistemic	5	37	42	27,68
Deontic	95	63	158	16,25
Dynamic	0	0	0	-
Total	100	100	200	43,93
$p=3,40 \times 10^{-11} < 0,05$				

Table 5. Modal verb *should* in Scottish English (SCOTS) and English English (BNC).

Deontic *should* appears most frequently in both corpora (95 of 100 in SCOTS and 63 of 100 in BNC):

(143) KP1 820 **Should** add it to dictionary DEONTIC

(144) HV2 202 I mean the army was left to do a job which politicians **should** have done. DEONTIC (BNC)

(145) F1145 I I only personally once skived and I was so worried about it that when my gran came in to visit my sister who was sick, I hid in the wardrobe, so I'd have been as well goin to school. It was only, it was the last day of term, and I thought, my sister was ill, it wasnae very fair that she wasnae goin to school and I **should** have //to go so// DEONTIC

(146) F812 //You **should** listen to what he's// what, some of the jobs he comes back, I, every summer and you go, "So Callum, what have you worked at this year?" he's like, "Well this year I was a removal man, then I worked in a library," or no, what was it you did, you were, //ehm, you've done// DEONTIC (SCOTS)

In SCOTS it has an obvious lead to the epistemic *should*, which appears only five times in total.

The distribution is more even in the BNC, where epistemic *should* appears 37 times:

(147) KE3 916 Oh [pause] wait a minute the doctor shouldn't have seen it then **should** he? EPISTEMIC

(148) KD6 4626 well I don't know how they work it, but I **shouldn't** get a bad one, it should be okay, I mean my dad's gonna give me money anyway, so, I mean if I, London City so Phil [gap:name] was saying one of the best in the country he says that, so hopefully if they, they give me an offer, the others should as well EPISTEMIC (BNC)

(149) F806 Tut, but my mum said that everything's all back in its place so that **should** be nice. I don't know what she meant by that, but. Yeah. EPISTEMIC

(150) F634 So, we just turned tail and went back, and we we sort of sat outside for a while thinkin, "Well, has he been yet? Has he been to the door yet? Or, or will he go after." Cause we thought, "Well, we need to go, we need to get some tea sometime", //So, we werenae sure whether [laugh] we **should** go// EPISTEMIC (SCOTS)

The findings hold with the statement of Carter and McCarthy (2006, 653), according to which *should* is most often used to express events that are desirable or considered to represent the ideal state of affairs.

The notion of the absence of *should* from Scots is taken up by Beal (1997, 366). If we consider the findings, this is not the case – *should* is very much present in Scottish English, although mainly in the deontic meaning. Miller (1993, 117) has in turn remarked that *should* is used to replace *ought* in Scots. This is also visible in the findings, as *ought to* appears only eight times in total in SCOTS. The difference between the distribution of *should* in SCOTS and BNC is statistically significant as the p value $3,40 \times 10^{-11}$ is smaller than 0,05.

5.6. Must

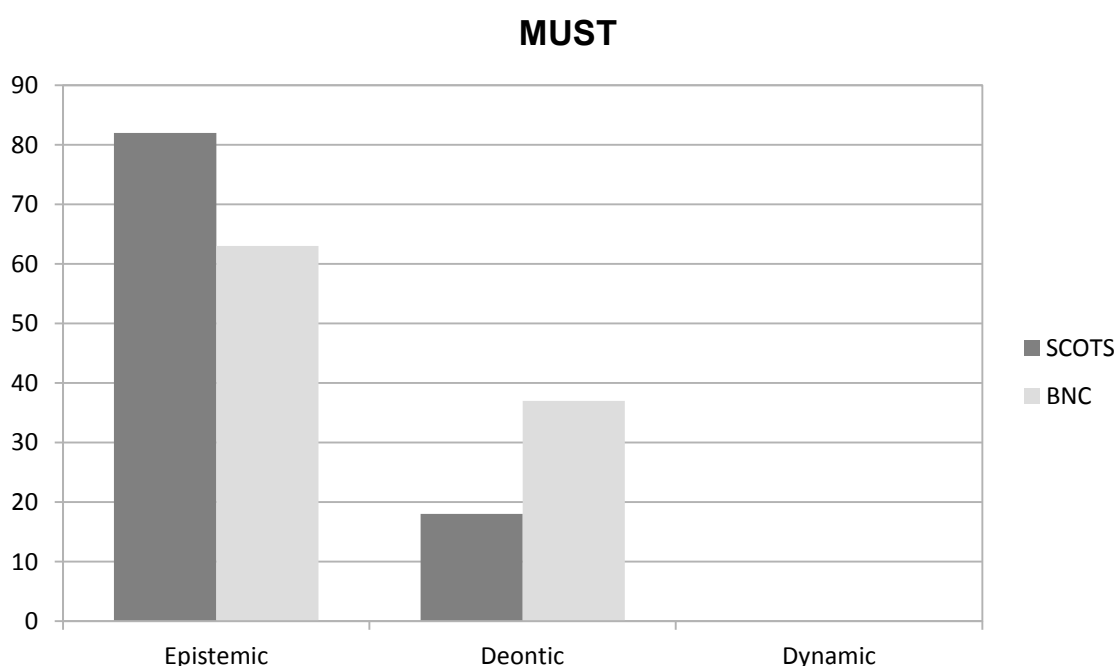


Figure 6. Modal verb *must* in Scottish English (SCOTS) and English English (BNC).

Must				
	SCOTS	BNC	Total	Chi-square
Epistemic	82	63	145	5,73
Deontic	18	37	55	9,76
Dynamic	0	0	0	-
Total	100	100	200	15,49
$p=8,31 \times 10^{-5} < 0,05$				

Table 6. Modal verb *must* in Scottish English (SCOTS) and English English (BNC).

As with *may*, there are no occurrences of dynamic *must* in either of the corpora, which can quite easily be explained by the notion that *must* is not used to express volition or ability to do something. The most common usage for *must* is expressing obligation (Carter and McCarthy 2006, 654-655). In addition to this it is often used to mark conclusion (ibid.). In Scottish English, *must* is reserved for expressing the latter meaning, and replaced by *have to* and *need to* when obligation is expressed (Miller 2004, 52).

As shown by the findings, deontic *must* is represented well enough in the BNC (37 units out of 100) but there are also occurrences of deontic *must* in SCOTS (18 out of 100) despite the more common usage of *have to* or *need to* in expressing obligation in *Scottish English*:

(151) KE6 5514 and somebody **must** try and get in touch with her in the evening if she's not in. DEONTIC

(152) KB0 1340 And I **must** say between us at the moment [pause] there wasn't a strong feeling [pause] what it's all about. DEONTIC (BNC)

(153) F1024 //not ankle biter, I **must** admit,// DEONTIC

(154) F1150 ...they are human beings and that society **must** be organized //in some way.// DEONTIC (SCOTS)

Epistemic *must* is well represented in both the corpora and especially SCOTS has many occurrences of it (SCOTS 82 out of 100, BNC 63 out of 100), which could be expected considering Miller's notion:

(155) KB1 1605 It **must** be the summer what does it. EPISTEMIC

(156) KCH 788 The paperwork **must** cost quite a lot [pause] that they keep pushing out. EPISTEMIC (BNC)

(157) F1149 //She mus- she **must** have been quite upset when she realised she was lost.// EPISTEMIC

(158) F1095 ... That's a sky. That's a red sky. It **must** be night-time. It's a red sky at... EPISTEMIC (SCOTS)

The difference in the distribution of *must* in Scottish English and English English is statistically significant with a p value of $8,31 \times 10^{-5}$.

5.7. Ought to

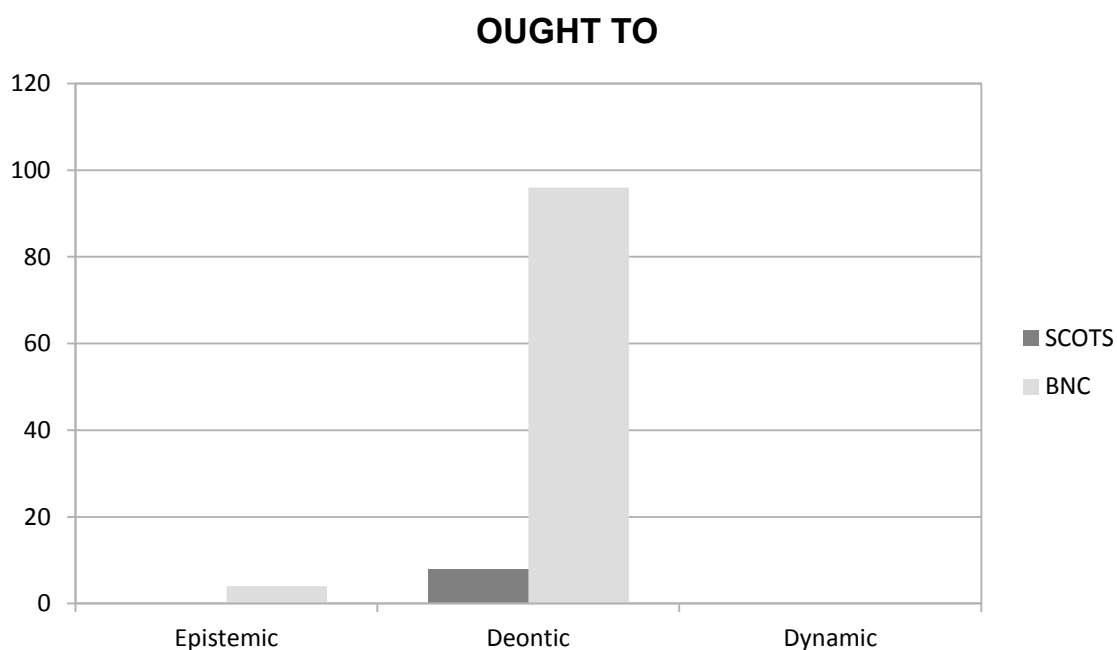


Figure 7. Semi-modal *ought to* in Scottish English (SCOTS) and English English (BNC).

Ought to				
	SCOTS	BNC	Total	Chi-square
Epistemic	0	4	4	4
Deontic	8	96	104	80,67
Dynamic	0	0	0	-
Total	8	100	108	84,67
$p=3,53 \times 10^{-20} < 0,05$				

Table 7. Semi-modal *ought to* in Scottish English (SCOTS) and English English (BNC).

According to Carter and McCarthy (2006, 659) *ought to* expresses states of affairs considered ideal or desired or something that is likely to happen. Of these usages the first mentioned, deontic modality, is well represented in the findings in the BNC (96 units out of 100), while the latter, epistemic modality, can be found only barely (4 occurrences in 100 units):

(159) KDW 6373 change the way that we thought children **ought to** be and educa--
DEONTIC

(160) KC0 2754 all that money that they keep demanding from the er th-- wa-- th--
wa-- the firms and such like that [pause] **ought to** have been ploughed back into the
firm! DEONTIC

(161) KBW 17869 So it **ought to** work out twelve and a six month [unclear]
EPISTEMIC

(162) KCS 1281 It **ought to** be it's his house hadn't he? EPISTEMIC (BNC)

Like *may*, also *ought to* should be nonexistent in Scottish English according to Miller (2004, 52). It does still appear in SCOTS, even though there are only eight occurrences in total. All of these represent deontic *ought to*:

(163) F1009 ...I work in Gaelic that they **ought to** put cheers in//in Gaelic as well.//
DEONTIC

(164) F1189 //[exhale]// //Well they **ought to**! [laugh]// DEONTIC

(165) F1093 I dinna think that goes in there. I think we **ought to** put this bit down the bottom, okay? Turn it round that way. DEONTIC

(166) F1093 Well I think it's bath time next. Mm? You **ought to** clean your seat after your tea. DEONTIC (SCOTS)

SCOTS did not have any occurrences of epistemic modality as compared to the four occurrences in the BNC, and neither of the corpora had any occurrences of dynamic *ought to*. The difference between Scottish English and English English was found to be statistically significant in the case of *ought to*, p value being $3,53 \times 10^{-20}$.

5.8. Have to

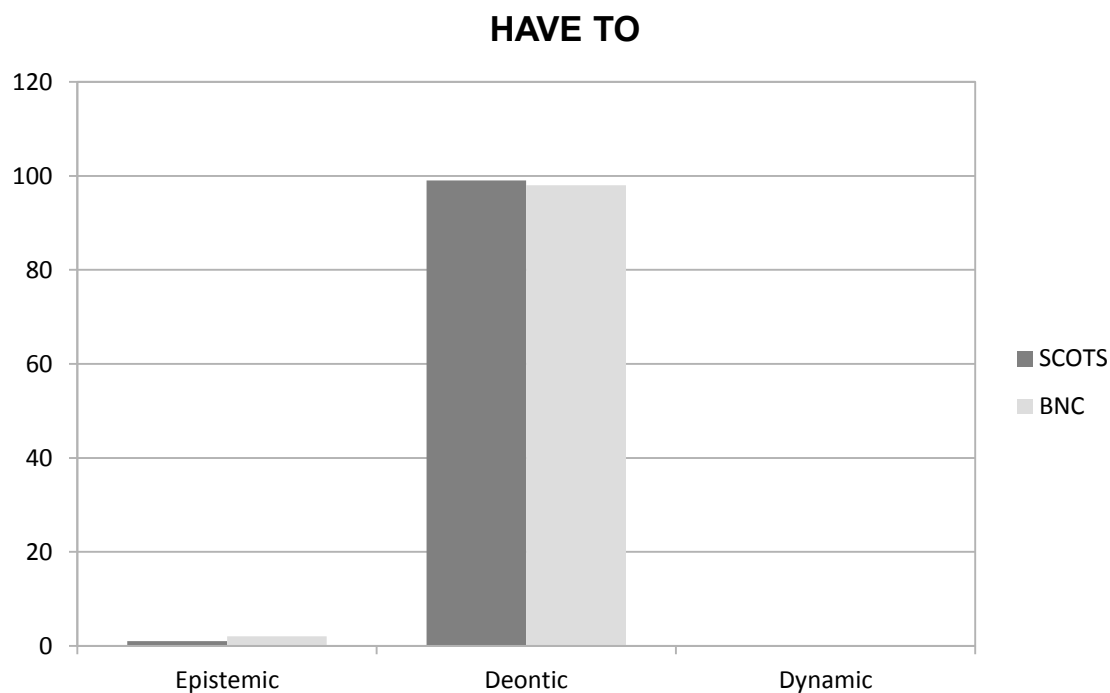


Figure 8. Semi-modal *have to* in Scottish English (SCOTS) and English English (BNC).

Have to				
	SCOTS	BNC	Total	Chi-square
Epistemic	1	2	3	0,5
Deontic	99	98	197	0,01
Dynamic	0	0	100	-
Total	100	100	200	0,51
p=0,48 > 0,05				

Table 8. Semi-modal *have to* in Scottish English (SCOTS) and English English (BNC).

Carter and McCarthy (2006, 667) note that *have to* is similar to *must* in its meanings – it is used to express obligation and conclusions. Considering that *have to* is said to refer to conclusions, epistemic *have to* appears only three times in total in the data – once in SCOTS and twice in the BNC:

(167) KCG 1436 Hopefully we can get away with that room Jane with just maybe
having to emulsion it EPISTEMIC

(168) KCK 329 It'll **have to** be dark. EPISTEMIC (BNC)

(169) M1108 Aye. And, and, Postman has to say, it **has to** be one hundred thousand
[inaudible] so Jessie and Pat can wake up. EPISTEMIC (SCOTS)

The rest of the occurrences represent deontic modality, thus following the notion made by Carter and McCarthy:

(170) KBW 15497 Wonder if we've got any milk to give Morgan, oh no we haven't
aha, when I take you home we'll **have to** get some more milk, otherwise we'll be
stranded tomorrow [pause] DEONTIC

(171) JYL 17 Er the problem with the job is it's very tight inside there so we've **had to**
use cranes and basically poured concrete [unclear] er [unclear] ours is being done by a
separate sub-contractor. DEONTIC (BNC)

(172) F1024 although the women **have to** wear hats in the court. Ehm even in, in
Dundee when I first joined, if you were in the C.I.D. plain clothes you had to wear a

hat if you were in court. Ehm but it all changed. Very much so, it's far more practical now to wear boots an a pair of trousers, especially if your havin tae //wrestle with folk or climb walls an deal with things,// DEONTIC
 (173) F943 But exactly, you go- you **have to** do it now. You've got to have written consent for everything. DEONTIC (SCOTS)

The distribution between the three types is almost identical in the two corpora and the difference was shown to lack statistical significance with a p value of 0,48.

Considering Scots, according to Miller (2004, 53) *have to* replaces *must* in Scots when obligation is expressed. Findings would seem to support this, as deontic *must* appears only 18 times in the data, as opposed to the 99 occurrences of deontic *have to*.

5.9. Conclusion

Considering Miller's (2004, 52) notion of the *Scottish English* modal system being very different from that of *Standard English*, the division of occurrences of the three types of modality is surprisingly similar with *may*. Miller (1993, 116-117) notes that *can*, *get to* or *get + gerund* are used instead of *may* when expressing permission. In light of the results this would seem to hold, as *may* appears only thrice in the deontic meaning in the sample from SCOTS whereas the amount of *can* in the same meaning is notably larger, 61 occurrences. The results show, however, that this does not differ very much from the sample drawn from the BNC, as deontic *may* only appears 17 times as compared to the clearly much larger occurrence of epistemic *may*, with 83 units. This finding is consistent with Beal's (1997, 367) remark of the decreasing use of *may* in more informal Standard English to express permission, which means that the usage of *may* does not differ drastically between Scottish English and English English.

Miller (2004, 52) has also stated that the absence of *may*, *ought* and *shall* is a feature considered to be distinctively Scottish. As seen in the discussion above, this definitely is not the case with *may* which is very much present in the current Scottish English with a distribution quite

similar to that of English English with epistemic modality being clearly the most often occurring form of *may* in both the corpora. Of the supposedly absent verbs in Scottish English *ought* is also present, although *ought to* does appear only eight times in total in SCOTS, which also makes the comparison with the BNC results on *ought to* slightly biased.

Can is noted to be most often used to express permission and ability in Standard English (Carter and McCarthy 2006, 642). The results from the BNC would seem to support this, as deontic and dynamic meaning are the two most common senses for *can* to appear in the BNC.

In Scots, *must* is reserved to express only conclusion, whereas in Standard English the area of usage is conclusion and obligation (Miller 2004, 52). Despite the fact that *must* does appear the most in the epistemic sense in SCOTS (82 occurrences), it does still also appear in the deontic sense, which in fact expresses obligation among its other meanings, exemplified by 18 occurrences. The pattern is the same for English English in the BNC, with 63 occurrences of epistemic *must* and 37 occurrences of deontic *must*. The use of *must* would not seem to be that different in Scottish English as compared with English English so as to make the way it is used in Scottish English distinctively Scottish.

Naturally this still leaves a lot of space for differences but shows that the modal systems of *Scottish English* and *Standard English* are not that different all the time and when they are, it is not necessarily that obvious which one has taken the step towards the other. A striking notion during the research was the absence of double modals in the sample taken from SCOTS, despite double modals being one of the most notable differentiating features of *Scottish English* (Miller 2004, 53).

Bibliography

Corpora

SCOTS = Scottish Corpus of Texts and Speech. [Internet] Available from

<http://www.scottishcorpus.ac.uk/>.

BNC = The British National Corpus. [Internet] Available from <http://www.natcorp.ox.ac.uk/>.

Other sources

Aarts, Bas, Sylvia Chalker and Edmund Weiner. 2014. *The Oxford Dictionary of English Grammar*.

Oxford: Oxford University Press.

Anderson, Wendy and John Corbett. 2009. *Exploring English with Online Corpora: An*

Introduction. New York: Palgrave Macmillan.

Beal, Joan. 1997. "Syntax and Morphology." In *The Edinburgh History of the Scots Language*, ed.

Charles Jones, 335-377. Edinburgh: Edinburgh University Press.

Biber, Douglas, Stig Johansson, Geoffrey Leech, Susan Conrad and Edward Finegan. 1999.

Longman Grammar of Spoken and Written English. Harlow: Pearson Education Limited.

Carter, Ronald and Michael McCarthy. 2006. *Cambridge Grammar of English: A Comprehensive*

Guide: Spoken and Written English: Grammar and Usage. Cambridge: Cambridge

University Press.

Chambers, J.K. and Peter Trudgill. 1998. *Dialectology*. Cambridge: Cambridge University Press.

Collins, Peter. 2008. "The English modals and semi-modals: Regional and stylistic variation." In

The Dynamics of Linguistic Variation: Corpus evidence on English past and present, ed.

Terttu Nevalainen, Irma Taavitsainen, Päivi Pahta and Minna Korhonen, 129-145.

Amsterdam: John Benjamins Publishing Company.

- Douglas, Fiona. 2006. "English in Scotland." In *The Handbook of World Englishes*, ed. Braj B. Kachru, Yamuna Kachru and Cecil L. Nelson, 41-57. Oxford: Blackwell Publishing Ltd.
- Douglas, Fiona. 2009. *Scottish Newspapers, Language and Identity*. Edinburgh: Edinburgh University Press.
- Grant, William, ed. 1931. *The Scottish National Dictionary, Volume 1*. Edinburgh: The Scottish National Dictionary Association.
- Grant, William and James Main Dixon. 1921. *Manual of Modern Scots*. Cambridge: Cambridge University Press.
- Huddleston, Rodney and Geoffrey K. Pullum. 2002. *Cambridge Grammar of the English Language*. Cambridge: Cambridge University Press.
- Lavrakas, Paul J., ed. 2008. *Encyclopedia of Survey Research Methods*. California: SAGE Publications.
- Leech, Geoffrey, Marianne Hundt, Christian Mair and Nicholas Smith. 2009. *Change in Contemporary English: A Grammatical Study*. Cambridge: Cambridge University Press.
- Luojola, Teemu. 2006. Kielitieteellisen aineiston kvantitatiiviset analyysimenetelmät. [Internet] University of Helsinki. Available from http://www.ling.helsinki.fi/~fkarlsso/methods/kvant_men.pdf. [Accessed 21 April 2015]
- Macafee, Caroline and A.J. Aitken. 2002. "A history of Scots to 1700." In *A Dictionary of the Older Scottish Tongue* vol. XII, xxix-clvii. [Internet] Available from <http://www.dsl.ac.uk/about-scots/history-of-scots/origins/>. [Accessed 26 April 2015]
- Mather, J.Y. and H.H. Speitel, eds. 1975. *The Linguistic Atlas of Scotland: Scots section, Volume 1*. London: Croom Helm.
- McArthur, Tom. 1979. "The Status of English in and furth of Scotland." In *Languages of Scotland*, ed. A.J. Aitken and Tom McArthur, 50-67. Edinburgh: Chambers.
- McArthur, Tom. 2003. *The English Languages*. Cambridge: Cambridge University Press.

- McCarthy, Michael and Anne O’Keeffe. 2009. “Corpora and spoken language.” In *Corpus Linguistics: An International Handbook, Volume 2*, ed. Anke Lüdeling and Merja Kytö, 1008-1024. Berlin: Mouton de Gruyter.
- McClure, J. Derrick. 1994. “English in Scotland.” In *The Cambridge History of the English Language, Volume 5: English in Britain and Overseas, Origins and Development*, ed. Robert Burchfield, 23-93. Cambridge: Cambridge University Press.
- Meyer, Charles F. 2002. *English Corpus Linguistics: An Introduction*. Cambridge: Cambridge University Press.
- Miller, Jim. 1993. “The Grammar of Scottish English.” In *Real English: The Grammar of English Dialects in the British Isles*, ed. James Milroy and Leslie Milroy, 99-138. London: Longman.
- Miller, Jim. 2004. “Scottish English: morphology and syntax.” In *A Handbook of Varieties of English, Volume 2: Morphology and Syntax*, ed. Bernd Kortmann, Kate Burridge, Rajend Mesthrie, Edgar W. Schneider and Clive Upton, 47-72. Berlin: Mouton de Gruyter.
- Murison, David. 1979. “The Historical Background.” In *Languages of Scotland*, ed. A.J. Aitken and Tom McArthur, 2-13. Edinburgh: Chambers.
- Newbrook, Mark. 2002. “Dialectology.” In *The Linguistics Encyclopedia*, ed. Kirsten Malmkjær, 108-114. New York: Routledge.
- Nuzzo, Regina. 2014. “Statistical Errors.” *Nature* 506: 150-152.
- Palmer, F.R. 1979. *Modality and the English Modals*. London: Longman.
- Palmer, F.R. 1986. *Mood and modality*. Cambridge: Cambridge University Press.
- Palmer, Frank. 2003. “Modality in English: Theoretical, descriptive and typological issues.” In *Modality in Contemporary English*, ed. Roberta Facchinetti, Manfred Krug and Frank Palmer, 1-17. Berlin: Mouton de Gruyter.

- Rajalahti, Kaisa. 2006. The Modal and Quasi-Modal Verbs of Obligation and Necessity in Philippine and Singapore English. [Internet] Available from <http://urn.fi/urn:nbn:fi:uta-1-16247>. [Accessed 4 November 2015]
- Romaine, S. 1994. "Dialect and Dialectology." In *The Encyclopedia of Language and Linguistics, Volume 2*, ed. R.E. Asher and J.M.Y Simpson, 900-907. Oxford: Pergamon Press.
- Ruohonen, Keijo. 2011. Tilastomatematiikka. [Internet] Available from <http://math.tut.fi/~ruohonen/TM.pdf>. [Accessed 26 April 2015]
- Schmied, Josef. 2009. "Contrastive corpus studies." In *Corpus Linguistics: An International Handbook, Volume 2*, ed. Anke Lüdeling and Merja Kytö, 1140-1159. Berlin: Mouton de Gruyter.
- Swan, Michael. 2005. *Practical English Usage*. Oxford: Oxford University Press.
- Vasko, Anna-Liisa. 2010. "Background: Regional English Speech Project and the Cambridgeshire Dialect Grammar". *Cambridgeshire Dialect Grammar* (Studies in Variation, Contacts and Change in English 4). Helsinki: VARIENG.
<http://www.helsinki.fi/varieng/series/volumes/04/background.html>
- Viereck, Wolfgang. 2005. "Dialectology." In *Encyclopedia of Linguistics, Volume 1*, ed. Philipp Strazny, 266-268. New York: Fitzroy Newborn.
- Yule, George. 2006. *The Study of Language*. Cambridge: Cambridge University Press.